

TRANSMITTAL LETTER TO THE UNITED STATES
DESIGNATED/ELECTED OFFICE (DO/EO/US)
CONCERNING A FILING UNDER 35 U.S.C. 371

P25,762 USA

U.S. APPLICATION NO. (IF KNOWN, SEE 37 CFR 1.5)

10/069304

INTERNATIONAL APPLICATION NO.
PCT/CA00/00966INTERNATIONAL FILING DATE
18 August 2000PRIORITY DATE CLAIMED
19 August 1999 and 13 October 1999

TITLE OF INVENTION

Proline-rich Extensin-like Receptor Kinases

APPLICANT(S) FOR DO/EO/US

Daphne Goring and Nancy Silva

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☒ This is an express request to begin national examination procedures (35 U.S.C. 371(f)). The submission must include items (5), (6), (9) and (24) indicated below
4. ☒ The US has been elected by the expiration of 19 months from the priority date (Article 31).
5. ☒ A copy of the International Application as filed (35 U.S.C. 371 (c) (2))
 - a. ☐ is attached hereto (required only if not communicated by the International Bureau)
 - b. ☒ has been communicated by the International Bureau
 - c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☐ An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).
 - a. ☐ is attached hereto
 - b. ☐ has been previously submitted under 35 U.S.C. 154(d)(4).
7. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371 (c)(3))
 - a. ☐ are attached hereto (required only if not communicated by the International Bureau).
 - b. ☐ have been communicated by the International Bureau
 - c. ☐ have not been made, however, the time limit for making such amendments has NOT expired.
 - d. ☒ have not been made and will not be made
8. ☐ An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ☒ An oath or declaration of the inventor(s) (35 U.S.C. 371 (c)(4))
10. ☐ An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371 (c)(5))
11. ☐ A copy of the International Preliminary Examination Report (PCT/IPEA/409)
12. ☒ A copy of the International Search Report (PCT/ISA/210)

Items 13 to 20 below concern document(s) or information included:

13. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98
14. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included
15. ☒ A **FIRST** preliminary amendment
16. ☒ A **SECOND** or **SUBSEQUENT** preliminary amendment
17. ☐ A substitute specification
18. ☐ A change of power of attorney and/or address letter
19. ☐ A computer-readable form of the sequence listing in accordance with PCT Rule 13ter 2 and 35 U.S.C. 1.821 - 1.825
20. ☐ A second copy of the published international application under 35 U.S.C. 154(d)(4)
21. ☐ A second copy of the English language translation of the international application under 35 U.S.C. 154(d)(4)
22. ☒ Certificate of Mailing by Express Mail
23. ☒ Other items or information

Cover Page of International Publication No. WO 01/ 14563 A1

3. APPLICATION NO. (IF KNOWN, SEE 37 CFR 1.5)

10/069304

INTERNATIONAL APPLICATION NO.

PCT/CA00/00966

ATTORNEY'S DOCKET NUMBER

P25,762 USA

24. The following fees are submitted

BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)) :

- ☐ Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO ... \$1040.00
- ☒ International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO \$890.00
- ☐ International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO \$740.00
- ☐ International preliminary examination fee (37 CFR 1.482) paid to USPTO but all claims did not satisfy provisions of PCT Article 33(1)-(4) \$710.00
- ☐ International preliminary examination fee (37 CFR 1.482) paid to USPTO and all claims satisfied provisions of PCT Article 33(1)-(4) \$100.00

ENTER APPROPRIATE BASIC FEE AMOUNT =

\$890.00

Surcharge of \$130.00 for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492 (e))

\$0.00

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE
Total claims	20 - 20 =	0	x \$18.00
Independent claims	3 - 3 =	0	x \$84.00

\$0.00

\$0.00

Multiple Dependent Claims (check if applicable) ☐

\$0.00

TOTAL OF ABOVE CALCULATIONS =

\$890.00

☒ Applicant claims small entity status See 37 CFR 1.27) The fees indicated above are reduced by 1/2

\$445.00

SUBTOTAL =

\$445.00

Processing fee of \$130.00 for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492 (f))

\$0.00

TOTAL NATIONAL FEE =

\$445.00

Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31) (check if applicable). ☐

\$0.00

TOTAL FEES ENCLOSED =

\$445.00

Amount to be:
refunded \$
charged \$

- a. ☒ A check in the amount of 445.00 to cover the above fees is enclosed
- b. ☐ Please charge my Deposit Account No _____ in the amount of _____ to cover the above fees.
A duplicate copy of this sheet is enclosed
- c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No 19-5425 A duplicate copy of this sheet is enclosed
- d. ☐ Fees are to be charged to a credit card **WARNING:** Information on this form may become public **Credit card information should not be included on this form.** Provide credit card information and authorization on PTO-2038

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status

SEND ALL CORRESPONDENCE TO

Gene J. Yao
Synnestvedt & Lechner LLP
2600 Aramark Tower
1101 Market Street
Philadelphia, PA 19107-2950
Telephone: 215-923-4466
Facsimile: 215-923-2189

SIGNATURE

Gene J. Yao

NAME

47,193

REGISTRATION NUMBER

February 19, 2002

DATE

LAW OFFICES OF
SYNNESTVEDT & LECHNER LLP

2600 ARAMARK TOWER
1101 MARKET STREET
PHILADELPHIA, PA 19107-2950
TELEPHONE (215) 923-4466
FACSIMILE (215) 923-2189
E-MAIL synnlech@synnlech.com
www.synnlech.com

February 19, 2002

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of D. Goring and N. Silva
U.S. Application No. Unassigned
Filed Concurrently Herewith
Proline-rich Extensin-like Receptor Kinases

(Atty. Docket No. P 25,762 USA)

CERTIFICATE OF EXPRESS MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service, "Express Mail Post Office to Addressee", Mailing Label No. EL 930922731 US addressed to: Commissioner for Patents, Box Patent Application, Washington, D.C. 20231 on Tuesday, February 19, 2002.


Jaclyn M. Schlitter

Commissioner for Patents
Box Patent Application
Washington, D.C. 20231

PRELIMINARY AMENDMENT UNDER 37 CFR § 1.115
REDUCING THE NUMBER OF CLAIMS PRIOR TO CALCULATION OF THE FILING FEE

Sir:

Applicants request entry of the following amendments.

February 19, 2002
Page 2

In the Description

Page 1, after line 2, please insert the following paragraph.

--This is a continuation-in-part of international Application No.
PCT/CA00/00966, filed August 18, 2000, which claims priority to U.S. provisional
Application No. 60/149,466, filed August 19, 1999, and to U.S. provisional
Application No. 60/159,122, filed October 13, 1999.--

In the Claims

Please cancel Claims 15 to 18, 22, 24, 26, 27, and 29 to 47 without prejudice.

Respectfully submitted,



Gene J. Yao, Esquire
Reg. No. 47,193
Attorney for Applicants

Synnestvedt & Lechner LLP
Suite 2600
Aramark Tower
1101 Market Street
Philadelphia, PA 19107-2950
(215) 923-4466

GJY

10/069304

JC13 Rec'd PCT/PTO 19 FEB 2002

LAW OFFICES OF
SYNNESTVEDT & LECHNER LLP
2600 ARAMARK TOWER
1101 MARKET STREET
PHILADELPHIA, PA 19107-2950
TELEPHONE (215) 923-4466
FACSIMILE (215) 923-2189
E-MAIL synnlech@synnlech.com
www.synnlech.com

February 19, 2002

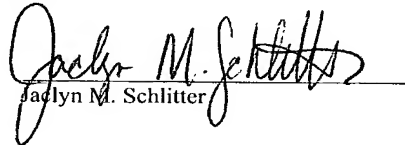
IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of D. Goring and N. Silva
U.S. Application No. Unassigned
Filed Concurrently Herewith
Proline-rich Extensin-like Receptor Kinases

(Atty. Docket No. P 25,762 USA)

CERTIFICATE OF EXPRESS MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service, "Express Mail Post Office to Addressee", Mailing Label No. EL 93092273 addressed to: Commissioner for Patents, Box Patent Application, Washington, D.C. 20231 on Tuesday, February 19, 2002.


Jaclyn M. Schlitter

Commissioner for Patents
Box Patent Application
Washington, D.C. 20231

SECOND PRELIMINARY AMENDMENT UNDER 37 CFR § 1.115

Sir:

Applicants request entry of the following amendments.

In the Claims

6. (Amended) An isolated nucleic acid molecule encoding a PERK polypeptide, a fragment of a PERK polypeptide having PERK activity, or a polypeptide having PERK activity, comprising a nucleic acid molecule selected from the group consisting of: (a) a nucleic acid molecule that hybridizes to a nucleic acid molecule consisting of [SEQ ID NO:1], or a complement thereof under low, moderate or high stringency hybridization conditions wherein the nucleic acid molecule encodes a PERK polypeptide or a polypeptide having PERK activity; and (b) a nucleic acid molecule degenerate with respect to (a), wherein the nucleic molecule encodes a PERK polypeptide or a polypeptide having PERK activity.
9. (Amended) The nucleic acid molecule of claim 1, wherein the PERK polypeptide comprises a PERK1 polypeptide.
14. (Amended) A recombinant nucleic acid molecule comprising a nucleic acid molecule of claim 1 and a constitutive promoter sequence or an inducible promoter sequence, operatively linked so that the promoter enhances transcription of the nucleic acid molecule in a host cell.
19. (Amended) A vector comprising the nucleic acid molecule of claim 1.
23. (Amended) A plant, a plant part, a seed, a plant cell or progeny thereof comprising the recombinant nucleic acid molecule of claim 14.
28. (Amended) An isolated polypeptide encoded by and/or produced from the nucleic acid molecule of claim 1.

February 19, 2002

Page 3

REMARKS

Claims 15 to 18, 22, 24, 26, 27, and 29 to 47 have been canceled without prejudice in a first Preliminary Amendment. Amendments have been made to Claims 9, 14, 19, and 28 to convert the claims from multiple-dependent form. An amendment of an editorial nature has been made to Claim 6. Support for the amendment to Claim 23 is found in the application at page 5, lines 28 and 29. No new matter has been added. Claims presently pending are Claims 1 to 14, 19 to 21, 23, 25, and 28.

A marked-up version of the claims amended, showing the changes made, is attached.

Respectfully submitted,



Gene J. Yao, Esquire
Reg. No. 47,193
Attorney for Applicants

Synnestvedt & Lechner LLP
Suite 2600
Aramark Tower
1101 Market Street
Philadelphia, PA 19107-2950
(215) 923-4466

VERSION WITH MARKINGS TO SHOW CHANGES MADE

6. (Amended) An isolated nucleic acid molecule encoding a PERK polypeptide, a fragment of a PERK polypeptide having PERK activity, or a polypeptide having PERK activity, comprising a nucleic acid molecule selected from the group consisting of: [(b)] (a) a nucleic acid molecule that hybridizes to a nucleic acid molecule consisting of [SEQ ID NO:1], or a complement thereof under low, moderate or high stringency hybridization conditions wherein the nucleic acid molecule encodes a PERK polypeptide or a polypeptide having PERK activity; and (b) a nucleic acid molecule degenerate with respect to (a), wherein the nucleic molecule encodes a PERK polypeptide or a polypeptide having PERK activity.
9. (Amended) The nucleic acid molecule of [any of claims 1 to 8] claim 1, wherein the PERK polypeptide comprises a PERK1 polypeptide.
14. (Amended) A recombinant nucleic acid molecule comprising a nucleic acid molecule of [any of claims 1 to 4] claim 1 and a constitutive promoter sequence or an inducible promoter sequence, operatively linked so that the promoter enhances transcription of the nucleic acid molecule in a host cell.
19. (Amended) A vector comprising the nucleic acid molecule of [any one of claims 1 to 4] claim 1.
21. (Amended) A host cell comprising the recombinant nucleic acid molecule of claim 1 [or the vector of claim 19], or progeny of the host cell.

23. (Amended) A plant, a plant part, a seed, a plant cell or progeny thereof comprising the recombinant nucleic acid molecule of [claim 15 or the vector of claim 19] claim 14.
28. (Amended) An isolated polypeptide encoded by and/or produced from the nucleic acid molecule of [any of claims 1 to 4, or the vector of claim 19] claim 1.



5000

DT11 Rec'd PCT/PTO JUN 19 2002

PET

#4

LAW OFFICES OF
SYNNESTVEDT & LECHNER LLP
2600 ARAMARK TOWER
1101 MARKET STREET
PHILADELPHIA, PA 19107-2950
TELEPHONE (215) 923-4466
FACSIMILE (215) 923-2189
E-MAIL synnlech@synnlech.com
www.synnlech.com

June 12, 2002

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

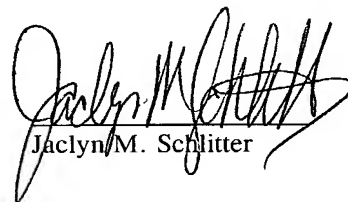
In re application of D. Goring and N. Silva
U.S. Application No. 10/069,304
Filed February 19, 2002
Proline-rich Extensin-like Receptor Kinases

Group No. NYA
Examiner NYA

(Atty. Docket No. P 25,762 USA)

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to the Commissioner for Patents, Washington, D.C. 20231, on Wednesday, June 12, 2002.


Jaclyn M. Schlitter

Commissioner for Patents
Washington, D.C. 20231

THIRD PRELIMINARY AMENDMENT UNDER 37 CFR § 1.115

Sir:

Applicants request entry of the following amendment under 37 CFR §1.115(b)(2). The status of the present application is that applicants are awaiting

U.S. Application No. 10/069,304

June 12, 2002
Attorney Docket No. P 25,762 USA
Page 2

receipt of a first Action on-the-merits.

Applicants note that a copy of the present Amendment was originally filed on February 22, 2002. However, according to the PTO, that copy was not successfully matched with the present application. Applicants request, therefore, that the present copy serve as the official copy of the present Amendment and that the PTO ignore the copy mailed February 22, 2002.

In the Description

Please delete the paragraph added by the first Preliminary Amendment, filed February 19, 2002, (if already added to the specification, the paragraph starting at page 1, line 3) and add the following paragraph before the heading "Field of the Invention".

--This is a U.S. national stage application based on international Application No. PCT/CA00/00966, filed August 18, 2000, which claims priority to U.S. provisional Application No. 60/149,466, filed August 19, 1999, and to U.S. provisional Application No. 60/159,122, filed October 13, 1999.--

REMARKS

The above amendment has been made to correct a clerical error made in a first Preliminary Amendment, filed February 19, 2002. In that Amendment, the present application was misidentified as being a continuation-in-part of international Application No. PCT/CA00/00966 when, in fact, it is a national stage application based thereon with no new matter added. Applicants confirm that no new matter

SYNNESTVEDT & LECHNER LLP

U.S. Application No. 10/069,304

June 12, 2002
Attorney Docket No. P 25,762 USA
Page 3

has been added by the first and second Preliminary Amendments which were filed on February 19, 2002. No new matter has been added by the present amendment.

Respectfully submitted,



Gene J. Yao
Reg. No. 47,193
Attorney for Applicants

Synnestvedt & Lechner LLP
Suite 2600
Aramark Tower
1101 Market Street
Philadelphia, PA 19107-2950
(215) 923-4466



PTO/PGT Rec'd 06 AUG 2002

#3

July 29, 2002

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

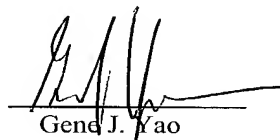
In re application of D. Goring and N. Silva
U.S. Application No. 10/069,304
Filed February 19, 2002
Proline-rich Extensin-like Receptor Kinases

Group No. NYA
Examiner NYA

(Atty. Docket No. P 25,762 USA)

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to the Commissioner for Patents, Washington, D.C. 20231, on Monday, July 29, 2002.



Gen J. Yao

Commissioner for Patents
Washington, D.C. 20231

FOURTH PRELIMINARY AMENDMENT UNDER 37 CFR § 1.115

Sir:

Applicants request entry of the following amendments.

In the Description

Please amend the paragraph, commencing at page 13, line 14, as follows.

Figure 11. Shows the nucleic acid molecules of [SEQ ID NOS:6 & 7] and the amino acid sequence of [SEQ ID NO:8].

In a preferred embodiment, the figure shows the sequence of the predicted Arabidopsis gene - Accession number AAC98010

- A) Genomic Sequence [SEQ ID NO:6]. The predicted open reading frame is underlined. The start codon (ATG) and stop codon (TGA) are double underlined.
- B) Translation of the predicted open reading frame [SEQ ID NOS:7 & 8]. The transmembrane domain is underlined.

Please amend the paragraph, commencing at page 13, line 22, as follows.

Figure 12. Shows the nucleic acid molecules of [SEQ ID NOS:9 & 10] and the amino acid sequence of [SEQ ID NO:11].

In a preferred embodiment, the figure shows the sequence of the predicted Arabidopsis gene - Accession number AAD15491

- A) Genomic Sequence [SEQ ID NO:9]. The predicted open reading frame is underlined. The start codon (ATG) and stop codon (TGA) are double underlined.
- B) Translation of the predicted open reading frame [SEQ ID NOS:10 & 11]. The transmembrane domain is underlined.

Please amend the paragraph, commencing at page 13, line 30, as follows.

Figure 13. Shows the nucleic acid molecules of [SEQ ID NOS:12 & 13] and the amino acid sequence of [SEQ ID NO:14].

In a preferred embodiment, the figure shows the sequence of the predicted

Arabidopsis gene - Accession number CAA18823.

- A) Genomic Sequence [SEQ ID NO:12]. The predicted open reading frame is underlined. The start codon (ATG) and stop codon (TGA) are double underlined.
- B) Translation of the predicted open reading frame [SEQ ID NOS:13 & 14]. The transmembrane domain is underlined.

Please amend the paragraph, commencing at page 14, line 7, as follows.

Figure 14. Shows the nucleic acid molecules of [SEQ ID NOS:15 & 16] and the amino acid sequence of [SEQ ID NO:17].

In a preferred embodiment, the figure shows the sequence of the predicted Arabidopsis gene - Accession number CAA18590

- A) Genomic Sequence [SEQ ID NO:15]. The predicted open reading frame is underlined. The start codon (ATG) and stop codon (TGA) are double underlined.
- B) Translation of the predicted open reading frame [SEQ ID NOS: 15 & 16]. The transmembrane domain is underlined.

Please amend the paragraph, commencing at page 43, line 12, as follows.

The isolation of novel *Brassica napus* receptor kinases relied upon the newly constructed cDNA library and involved *in vivo* mass excision of the pBluecsript phagemids from the Uni-ZAP XR vectors as outlined by the manufacturer (Stratagene, La Jolla, CA). Following efficient mass excision, phagemid DNA was extracted using a large scale alkaline protocol as described by Sambrook et al. (1989) and subjected to the polymerase chain reaction (PCR) using two separate oligonucleotide combinations, RK1/RK2 and RK1/RK3 (obtained from M. Cock, École Normale Supérieure de Lyon, France) specifically designed to prime conserved subdomains of

the catalytic domain of receptor protein kinases. RK1 (5'-ggiggTTTCggiAT^T_CgTiTT^A_TAA^A_ggg - 3'; [SEQ ID NO:18]) served as the forward primer and was constructed based upon a conserved amino acid consensus (GGFGIV^F/_YKG; [SEQ ID NO:19]) within subdomain I of the catalytic domain. The degeneracy of one reverse primer RK2 (5' - AAiATiC^T_gigCCATiCC^A_gAA^A_gTC - 3'; [SEQ ID NO:20]) reflects a conserved amino acid consensus (DFGMARIF; [SEQ ID NO:21]) of subdomain VII which closely resembles the SRKs in Brassica. The second reverse oligonucleotide RK3 (5' - A^g_AiA^g_ATC^TTTigCiA^A_giCC^A_gAA^A_gTC - 3'; [SEQ ID NO:23]) was generated based upon conserved amino acids (DFGLAKLL; [SEQ ID NO:24]) within subdomain VII prevalent among the RLKs isolated in Arabidopsis. Phagemid DNA was amplified in a reaction mixture containing 1 microliter of excised phagemid DNA, 10x PCR buffer (100mM Tris-HCl pH8.3, 500mM KCl, 15mM MgCl₂), 10mM deoxyribonucleotide triphosphate mixture, 10 micromolar of each oligonucleotide primer and 0.5 microliter Tsg polymerase (BioBasics, Canada). The PCR reaction was heated at 95°C for 2 min and amplified for 35 cycles under the following amplification conditions: 1 min at 95°C for denaturation, 1 min 30 sec at 50°C for primer annealing and 1 min at 72°C for synthesis. A final extension cycle of 10 min at 72°C was also incorporated into the amplification program. All PCR products generated of the expected size (420-450 bp) were gel purified, cloned into the pT7Blue plasmid (Novagen, Madison, WI) and introduced into *Escherichia coli* DH5- alpha. Transformants were tested for the presence of an insert and positive clones were sequenced with universal primers (R-20 and U-19) by an ABI automated sequencer (Model 373 STRETCH DNA; Perkin Elmer Corp., Canada Ltd.) using the dideoxychain-terminating method described by Sanger et al. (1977). Sequence analyses performed using DNAsis[®] software (Hitachi Software, San Bruno, CA) at the nucleotide and amino acid levels.

Please amend the paragraph, commencing at page 45, line 5, as follows.

The 5' end of the PERK1 cDNA was obtained by the procedure for the rapid amplification of cDNA ends originally described by Frohman et al. (1988) using the 5' RACE System, Version 2.0 kit (Gibco-BRL, Gaithersburg, MD). First strand cDNA was synthesized from approximately 300:g of mixed Westar and W1 pistil total RNA using a gene specific primer GSP1 (5'-TAACCAACAAGAgACA-3'; [SEQ ID NO:22]) designed to anneal approximately 300 bp from the 5' end of the PERK1 cDNA (1512 bp) isolated from the library screen. Following cDNA synthesis, the first strand product was purified from unincorporated dNTPs and GSP1 using a GLASS MAX[®] spin cartridge. A homopolymeric tail was added to the 3' end of the cDNA using TdT (terminal deoxynucleotidyl transferase) and dCTP. Tailed cDNA was amplified using a second gene specific primer GSP2 (5'-CCACTCCCAACTTTCAAC-3'; [SEQ ID NO:25]) designed to anneal 3' to GSP1 with respect to the cDNA, and an abridged anchor primer (Gibco-BRL, Gaithersburg, MD) which annealed to the homopolymeric tail. PCR amplification was carried out for 35 cycles of denaturation at 94°C for 1 min, primer annealing at 55°C for 1 min and extension at 72°C for 2 min, followed by a final extension cycle for 10 min. A PCR product of the expected size (~1 kb) corresponding to the 5' end of PERK1 was gel purified, cloned into the pT7Blue plasmid (Novagen, Madison, WI) and transformed into *Escherichia coli* DH5- alpha. Confirmation of the 5'RACE product was obtained by plasmid Southern blot analysis as described above and by sequential primer based sequencing.

The paragraph, commencing at page 45, line 25, has been amended as follows.

A PCR based approach was used to generate a full length PERK1 cDNA by combining the 5'RACE product cloned into the EcoRV site of pT7Blue with the cDNA isolated from the library screen cloned into the EcoRI/XhoI sites of the

pBluescript SK phagemid. A forward primer (5'-ggAAAgCTTgCATgCCTgCAggTCgAC -3'; [SEQ ID NO:26]) containing an internal PstI site was designed to anneal upstream to the EcoRV cloning site of pT7Blue. A reverse primer (5'-CgCCTgCAGgTAATACgACTCACTATAggg -3'; [SEQ ID NO:27]) also containing a PstI site was designed based on pBluescript phagemid sequence immediately 3' to the EcoRI/XhoI cloning site. Full length PERK1 cDNA was generated from a 100 microliter PCR reaction containing 1 microliter (~20ng) of each template (cDNA in pT7Blue and pBluescript phagemid), 10x Pfu Buffer (200mM Tris-HCl pH8.8, 100mM (NH₄)₂SO₄, 20mM MgSO₄, 1% Triton[®]X-100, 1mg/mlBSA), 10mM dNTPs, 50pmol forward and reverse primers and 1microliter Pfu polymerase (Gibco-BRL, Gaithersburg, MD). The samples were heated to 94°C for 5 min and amplified for 30 cycles with a denaturing cycle of 1 min, a primer annealing cycle at 53°C for 1 min followed by an extension cycle for 3 min at 72°C. The resulting PCR product of the expected size (~2.2kb) was gel purified and cloned into the PstI restriction site of pBluescript KS (+/-) II. The full length PERK1 cDNA sequence was confirmed by a sequential primer based sequencing approach using both universal and sequence specific primers as previously described. All DNA and protein sequence analysis was performed using the DNAsis[®] Software (Hitachi Software, San Bruno,CA).

In the Drawings

Please add the attached Figures 1(c), 1(d), and 1(e).

REMARKS

The status of the present application is that applicants have not yet received a first Action on-the-merits. Amendments have been made to the descriptive portion of

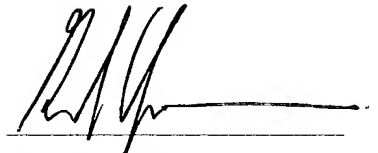
U.S. Application No. 10/069,304

July 29, 2002
Attorney Docket No. P 25,762 USA
Page 7

the specification to bring it in conformity with a Sequence Listing being submitted concurrently. In addition, Figures 1(c), 1(d), and 1(e) were added. These figures were inadvertently left out of the present application but were included with U.S. provisional Application No. 60/159,122 which is incorporated into the present application by reference (see page 48 of the present application). No new matter has been added by the present amendment.

A marked-up version of the changes made to the specification by the current amendment is attached.

Submitted respectfully,

A handwritten signature in black ink, appearing to read 'Gene J. Yao', written over a horizontal line.

Gene J. Yao
Reg. No. 47,193
Attorney for Applicants

Synnestvedt & Lechner LLP
Suite 2600, Aramark Tower
1101 Market Street
Philadelphia, PA 19107-2950
(215) 923-4466 - Telephone
(215) 923-2189 - Facsimile

VERSION WITH MARKINGS TO SHOW CHANGES MADE

In the Description

The paragraph, commencing at page 13, line 14, has been amended as follows.

Figure 11. Shows the nucleic acid molecules of [SEQ ID NOS:6 & 7] [[SEQ ID NO.: 6]] and the amino acid sequence of [SEQ ID NO:8] [[SEQ ID NO.: 7]].

In a preferred embodiment, the figure shows the sequence of the predicted Arabidopsis gene - Accession number AAC98010

- A) Genomic Sequence [SEQ ID NO:6]. The predicted open reading frame is underlined. The start codon (ATG) and stop codon (TGA) are double underlined.
- B) Translation of the predicted open reading frame [SEQ ID NOS:7 & 8]. The transmembrane domain is underlined.

The paragraph, commencing at page 13, line 22, has been amended as follows.

Figure 12. Shows the nucleic acid molecules of [SEQ ID NOS:9 & 10] [[SEQ ID NO.: 8]] and the amino acid sequence of [SEQ ID NO:11] [[SEQ ID NO.: 9]].

In a preferred embodiment, the figure shows the sequence of the predicted Arabidopsis gene - Accession number AAD15491

- A) Genomic Sequence [SEQ ID NO:9]. The predicted open reading frame is underlined. The start codon (ATG) and stop codon (TGA) are double underlined.

B) Translation of the predicted open reading frame [SEQ ID NOS:10 & 11]. The transmembrane domain is underlined.

The paragraph, commencing at page 13, line 30, has been amended as follows.

Figure 13. Shows the nucleic acid molecules of [SEQ ID NOS:12 & 13] [[SEQ ID NO.: 10]] and the amino acid sequence of [SEQ ID NO:14] [[SEQ ID NO.: 11]].

In a preferred embodiment, the figure shows the sequence of the predicted Arabidopsis gene - Accession number CAA18823.

C) Genomic Sequence [SEQ ID NO:12]. The predicted open reading frame is underlined. The start codon (ATG) and stop codon (TGA) are double underlined.

D) Translation of the predicted open reading frame [SEQ ID NOS:13 & 14]. The transmembrane domain is underlined.

The paragraph, commencing at page 14, line 7, has been amended as follows.

Figure 14. Shows the nucleic acid molecules of [SEQ ID NOS:15 & 16] [[SEQ ID NO.: 12]] and the amino acid sequence of [SEQ ID NO:17] [[SEQ ID NO.: 13]].

In a preferred embodiment, the figure shows the sequence of the predicted Arabidopsis gene - Accession number CAA18590

A) Genomic Sequence [SEQ ID NO:15]. The predicted open reading frame is underlined. The start codon (ATG) and stop codon (TGA) are double underlined.

B) Translation of the predicted open reading frame [SEQ ID NOS: 15 & 16]. The transmembrane domain is underlined.

The paragraph, commencing at page 43, line 12, has been amended as follows.

The isolation of novel *Brassica napus* receptor kinases relied upon the newly [constructedcDNA] constructed cDNA library and involved *in vivo* mass excision of the pBluecsript phagemids from the Uni-ZAP XR vectors as outlined by the manufacturer (Stratagene, La Jolla, CA). Following efficient mass excision, phagemid DNA was extracted using a large scale alkaline protocol as described by Sambrook et al. (1989) and subjected to the polymerase chain reaction (PCR) using two separate oligonucleotide combinations, RK1/RK2 and RK1/RK3 (obtained from M. Cock, École Normale Supérieure de Lyon, France) specifically designed to prime conserved subdomains of the catalytic domain of receptor protein kinases. RK1 (5-ggiggTTTCggiAT^Tc_AgTiTT^A_TC^TAA^A_ggg - 3'; [SEQ ID NO:18]) served as the forward primer and was constructed based upon a conserved amino acid consensus (GGFGIV^F/_YKG; [SEQ ID NO:19]) within subdomain I of the catalytic domain. The degeneracy of one reverse primer RK2 (5' - AAiATiC^T_gigCCATiCC^A_gAA^A_gC^T - 3'; [SEQ ID NO:20]) reflects a conserved amino acid consensus (DFGMARIF; [SEQ ID NO:21]) of subdomain VII which closely resembles the SRKs in Brassica. The second reverse oligonucleotide RK3 (5' - A^B_AiA^B_AC^TTTigCiA^A_giCC^A_gAA^A_gTC - 3'; [SEQ ID NO:23]) was generated based upon conserved amino acids (DFGLAKLL; [SEQ ID NO:24]) within subdomain VII prevalent among the RLKs isolated in Arabidopsis. Phagemid DNA was amplified in a reaction mixture containing 1 microliter of excised phagemid DNA, 10x PCR buffer (100mM Tris-HCl pH8.3, 500mM KCl, 15mM MgCl₂), 10mM deoxyribonucleotide triphosphate mixture, 10 micromolar of each oligonucleotide primer and 0.5 microliter Tsg polymerase (BioBasics, Canada). The PCR reaction was heated at 95°C for 2 min and amplified for 35 cycles under the following amplification conditions: 1 min at 95°C for denaturation, 1 min 30 sec at 50°C for primer annealing and 1 min at 72°C for

synthesis. A final extension cycle of 10 min at 72°C was also incorporated into the amplification program. All PCR products generated of the expected size (420-450 bp) were gel purified, cloned into the pT7Blue plasmid (Novagen, Madison, WI) and introduced into *Escherichia coli* DH5- alpha. Transformants were tested for the presence of an insert and positive clones were sequenced with universal primers (R-20 and U-19) by an ABI automated sequencer (Model 373 STRETCH DNA; Perkin Elmer Corp., Canada Ltd.) using the dideoxychain-terminating method described by Sanger et al. (1977). Sequence analyses performed using DNAsis® software (Hitachi Software, San Bruno, CA) at the nucleotide and amino acid levels.

The paragraph, commencing at page 45, line 5, has been amended as follows.

The 5' end of the PERK1 cDNA was obtained by the procedure for the rapid amplification of cDNA ends originally described by Frohman et al. (1988) using the 5' RACE System, Version 2.0 kit (Gibco-BRL, Gaithersburg, MD). First strand cDNA was synthesized from approximately 300:g of mixed Westar and W1 pistil total RNA using a gene specific primer GSP1 (5'-TAACCAACAAGACA-3'; [SEQ ID NO:22]) designed to anneal approximately 300 bp from the 5' end of the PERK1 cDNA (1512 bp) isolated from the library screen. Following cDNA synthesis, the first strand product was purified from unincorporated dNTPs and GSP1 using a GLASS MAX® spin cartridge. A homopolymeric tail was added to the 3' end of the cDNA using TdT (terminal deoxynucleotidyl transferase) and dCTP. Tailed cDNA was amplified using a second gene specific primer GSP2 (5'-CCACTCCCAACTTTCAAC-3'; [SEQ ID NO:25]) designed to anneal 3' to GSP1 with respect to the cDNA, and an abridged anchor primer (Gibco-BRL, Gaithersburg, MD) which annealed to the homopolymeric tail. PCR amplification was carried out for 35 cycles of denaturation

at 94°C for 1 min, primer annealing at 55°C for 1 min and extension at 72°C for 2 min, followed by a final extension cycle for 10 min. A PCR product of the expected size (~1 kb) corresponding to the 5' end of PERK1 was gel purified, cloned into the pT7Blue plasmid (Novagen, Madison, WI) and transformed into *Escherichia coli* DH5- alpha. Confirmation of the 5'RACE product was obtained by plasmid Southern blot analysis as described above and by sequential primer based sequencing.

The paragraph, commencing at page 45, line 25, has been amended as follows.

A PCR based approach was used to generate a full length PERK1 cDNA by combining the 5'RACE product cloned into the EcoRV site of pT7Blue with the cDNA isolated from the library screen cloned into the EcoRI/XhoI sites of the pBluescript SK phagemid. A forward primer (5'-ggAAAgCTTgCATgCCTgCAGgTCgAC -3'; [SEQ ID NO:26]) containing an internal PstI site was designed to anneal upstream to the EcoRV cloning site of pT7Blue. A reverse primer (5'-CgCCTgCAGgTAATACgACTCACTATAggg -3'; [SEQ ID NO:27]) also containing a PstI site was designed based on pBluescript phagemid sequence immediately 3' to the EcoRI/XhoI cloning site. Full length PERK1 cDNA was generated from a 100 microliter PCR reaction containing 1 microliter (~20ng) of each template (cDNA in pT7Blue and pBluescript phagemid), 10x Pfu Buffer (200mM Tris-HCl pH8.8, 100mM (NH₄)₂SO₄, 20mM MgSO₄, 1% Triton®X-100, 1mg/mlBSA), 10mM dNTPs, 50pmol forward and reverse primers and 1microliter Pfu polymerase (Gibco-BRL, Gaithersburg, MD). The samples were heated to 94°C for 5 min and amplified for 30 cycles with a denaturing cycle of 1 min, a primer annealing cycle at 53°C for 1 min followed by an extension cycle for 3 min at 72°C. The resulting PCR product of the expected size (~2.2kb) was gel purified and cloned into the PstI restriction site of pBluescript KS (+/-) II. The full length PERK1 cDNA

U.S. Application No. 10/069,304

July 29, 2002
Attorney Docket No. P 25,762 USA
Page 13

sequence was confirmed by a sequential primer based sequencing approach using both universal and sequence specific primers as previously described. All DNA and protein sequence analysis was performed using the DNAsis® Software (Hitachi Software, San Bruno,CA).

M:\GYao\25,000 to 25,999\25,762 USA\PTO\Fourth Preliminary Amendment wpd

-95 ttaactctctggtctccgtgtctcctctcttctcctgctgcttccttttaacactctctt -36
 -35 catttgccctttttgatttagatccaaagaagcagacATGTCCTCGGCGCCGTCTCCGGGG 24
 M S S A P S P G
 25 ACTGGTTCGCCTCCATCTCCACCATCAAACCTCCACAACCACCACTCCTCCTCCAGCTTCC 84
 T G S P P S P P S N* S T T T T P P P A S
 85 GCTCCTCCTCCCACCACACCTTCTTCTCCTCCGCGCCATCCACTATTCCGACATCTCCT 144
 A P P P T T P S S P P P P S T I P T S P
 145 CCTCCTTCTTCTCGCTCTACACCTTCTGCTCCTCCTCCATCTCCACCAACTCCATCTACG 204
 P P S S R S T P S A P P P S P P T P S T
 205 CCGGGATCTCCACCTCCTCTTCTCAGCCGTCTCCACCCGCTCCAACCTACGCCCGGATCT 264
 P G S P P P L P Q P S P P A P T T P G S
 265 CCACCCGCACCTGTTACTCCTCCTACTCGAAACCCTCCACCTTCAGTCCCAGGACCACCG 324
 P P A P V T P P T R N P P P S V P G P P
 325 TCCAATCCTTCACGCGAAGGAGGATCTCCTCGACCTCCATCTTCTCCCTCGCCGCCGTCT 384
 S N P S R E G G S P R P P S S P S P P S
 385 CCTTCTTCCGACGGTTTATCAACAGGAGTGGTGGTGGGAATCGCCATCGGAGGAGTCGCT 444
 P S S D G L S T G V V V G I A I G G V A
 445 CTGCTTGTGATAGTGACTCTGATTGTCTCCTCTGTAAGAAGAAACGACGGAGAGACGAA 504
 L L V I V T L I C L L C K K K R R R D E
 505 GAAGATGCTTACTATGTTCCCTCCGCCACCTCCTCCTGGTCCCAAAGCCGGAGGACCTTAC 564
 E D A Y Y V P P P P P P G P K A G G P Y
 565 GGTGGACAGCAGCAACAATGGCGGCAACAAAACGCAACACCACCGTCAGATCATGTCGTG 624
 G G Q Q Q Q W R Q Q N A T P P S D H V V
 625 ACGTCACTACCACCACCACCTAAGGCTCCATCTCCACCACGGCAACCTCCTCCACCTCCA 684
 T S L P P P P K A P S P P R Q P P P P P

FIG. 1C(1)

685 CCACCGCCTTTCATGAGCAGCAGCGGCGGCTCCGACTACTCGGACCGTCCAGTTCTTCCT 744
P P P F M S S S G G S D Y S D R P V L P

745 CCACCGTCTCCAGGGCTTGTGTTAGGCTTCTCCAAAAGCACTTTCACATACGAGGAGCTA 804
P P S P G L V L G F S K S T F T Y E E L

805 GCTAGAGCCACCAATGGTTTCTCCGAGGCGAACTTGTTAGGACAAGGCGGGTTTCGGTTAC 864
A R A T N ^IG F S E A N L L G Q G G F G Y

865 GTGCACAAAGGTGTGTTGCCTAGTGGGAAAGAAGTTGCTGTGAAGCAGTTGAAAGTTGGG 924
V H K G V L P S ^{II}G K E V A V K Q L K V G

925 AGTGCTCAGGGAGAGAGGGAGTTTTCAGGCAGAGGTTGAGATCATCAGCAGAGTTCACCAC 984
S G ^{III}Q G E R E F Q A E V E I I S R V ^{IV}H H

985 AGGCATCTGGTGTCTCTTGTGTTGGTTATTGCATCGCCGGTGCCAAAAGATTGCTTGTCTAT 1044
R H L V S L V G Y C I A ^VG A K R L L V Y

1045 GAGTTTGTTCCTAACAACAATCTCGAGCTTTCACCTCCATGGCGAGGGACGGCCTACAATG 1104
E F V P N N N L E L H L H G E G R ^{VI}P T M

1105 GAATGGAGCACCAGATTGAAGATTGCTCTTGGATCTGCTAAAGGACTTTCTTATCTTCAT 1164
E W S T R L K I A L G S A K G L S Y L H

1165 GAAGATTGCAATCCTAAAATCATTACCGTGATATCAAGGCTTCAAACATATTGATAGAT 1224
E D C N P K I I H R D I K A S N I L I D

1225 TTCAAGTTTGAAGCTAAGGTTGCTGATTTTGGTCTTGCTAAGATTGCTTCTGATACAAAC 1284
F K F ^{VII}E A K V A D F G L A K I A S D T N

1285 ACGCATGTATCAACACGTGTGATGGGAACCTTTGGGTACTTGGCTCCGGAATACGCTGCA 1344
T H V ^{VIII}S T R V M G T F G Y L A P E Y ^{IX}A A

1345 AGCGGAAAGCTCACGGAGAAGTCTGACGTTTTCTCATTTGGCGTTGTGCTTTTGGAGCTC 1404
S G K L T E K S D V F S F G V V L L E L

FIG. 1C(2)

1405 ATTACTGGACGTCGACCCGTTGATGCCAACAATGTCTATGTAGATGACAGCTTAGTTGAC 1464
I T G R R P V D A N N^XV Y V D D S L V D
 1465 TGGGCACGACCATTGCTTAACCGAGCATCTGAGCAAGGAGACTTTGAGGGTTTAGCTGAT 1524
W A R P L L N R A S E Q G D F E G L A D
 1525 GCAAAGATGAATAATGGGTATGACAGAGAGGAGATGGCTCGCATGGTTGCTTGTGCTGCG 1584
A K M N N G Y D R E^{XI}E M A R M V A C A A
 1585 GCTTGTGTTTCGCCATTTCAGCTCGCCGACACCTCGCATGAGCCAGATTGTGCGTGCCTTA 1644
A C V R M S A R R R P R M S Q I V R A L
 1645 GAAGGAAATGTATCACTGTCAGATCTTAACGAAGGGATGAGACCAGGTCAAAGCAATGTA 1704
 E G N* V S L S D L N E G M R P G Q S N V
 1705 TACAGCTCATACGGAGGAAGCACCGATTATGACTCGAGCCAGTACAATGAAGACATGAAG 1764
 Y S S Y G G S T D Y D S S Q Y N E D M K
 1765 AAGTTTAGGAAAATGGCACTTGGAACCTCAAGAGTACAACGCCACGGGTGAGTACAGTAAT 1824
 K F R K M A L G T Q E Y N* A T G E Y S N
 1825 CCGACCAGTGACTATGGACTGTACCCGTCTGGTTCAAGCAGCGAGGGCCAAACCACACGC 1884
 P T S D Y G L Y P S G S S S E G Q T T R
 1885 GAAATGGAGATGGGGAAGATTAAGAGAACCGGTCAGGGTTATAGTGACCTTCTCTTTAA 1944
 E M E M G K I K R T G Q G Y S G P S L *
 1945 accagatgggagagaaaattgaaggggtgtttttttcattattttttttaaaactgtaaagata 2004
 2005 tgagaaaaattgccttactctaattaaaaccactacgatataaggtttataatacgttttga 2064
 2065 attggtttttaaaaaaaaaaaaaaaaaaaaaa 2093

FIG. 1C(3)

10	20	30	40	50	60
TTAACTCTCT	GGTCTCCGTG	TCTCCTCTCT	TCTCCTGCTG	CTTCCTTTTA	ACACTCTCTT
70	80	90	100	110	120
CATTTGCCTT	TTTGATTAG	ATCCAAAGAA	GCAGACATGT	CCTCGGCGCC	GTCTCCGGGG
130	140	150	160	170	180
ACTGGTTCGC	CTCCATCTCC	ACCATCAAAC	TCCACAACCA	CCACTCCTCC	TCCAGCTTCC
190	200	210	220	230	240
GCTCCTCCTC	CCACCACACC	TTCTTCTCCT	CCGCCGCCAT	CCACTATTCC	GACATCTCCT
250	260	270	280	290	300
CCTCCTTCTT	CTCGCTCTAC	ACCTTCTGCT	CCTCCTCCAT	CTCCACCAAC	TCCATCTACG
310	320	330	340	350	360
CCGGGATCTC	CACCTCCTCT	TCCTCAGCCG	TCTCCACCCG	CTCCAACCTAC	GCCCCGGATCT
370	380	390	400	410	420
CCACCCGCAC	CTGTTACTCC	TCCTACTCGA	AACCTTCCAC	CTTCAGTCCC	AGGACCACCG
430	440	450	460	470	480
TCCAATCCTT	CACGCGAAGG	AGGATCTCCT	CGACCTCCAT	CTTCTCCCTC	GCCGCCGTCT
490	500	510	520	530	540
CCTTCTTCCG	ACGGTTTATC	AACAGGAGTG	GTGGTGGA	TCGCCATCGG	AGGAGTCGCT
550	560	570	580	590	600
CTGCTTGTGA	TAGTGACTCT	GATTTGTCTC	CTCTGTAAGA	AGAAACGACG	GAGAGACGAA
610	620	630	640	650	660
GAAGATGCTT	ACTATGTTCC	TCCGCCACCT	CCTCCTGGTC	CCAAAGCCGG	AGGACCTTAC
670	680	690	700	710	720
GGTGGACAGC	AGCAACAATG	GCGGCAACAA	AACGCAACAC	CACCGTCAGA	TCATGTCGTG
730	740	750	760	770	780
ACGTCACTAC	CACCACCACC	TAAGGCTCCA	TCTCCACCAC	GGCAACCTCC	TCCACCTCCA
790	800	810	820	830	840
CCACCGCCTT	TCATGAGCAG	CAGCGGCGGC	TCCGACTACT	CGGACCGTCC	AGTTCTTCTT
850	860	870	880	890	900
CCACCGTCTC	CAGGGCTTGT	GTTAGGCTTC	TCCAAAAGCA	CTTTCACATA	CGAGGAGCTA
910	920	930	940	950	960
GCTAGAGCCA	CCAATGGTTT	CTCCGAGGCG	AACTTGTTAG	GACAAGGCGG	GTTCGGTTAC
970	980	990	1000	1010	1020
GTGCACAAAG	GTGTGTTGCC	TAGTGGGAAA	GAAGTTGCTG	TGAAGCAGTT	GAAAGTTGGG
1030	1040	1050	1060	1070	1080
AGTGGTCAGG	GAGAGAGGGA	GTTTCAGGCA	GAGGTTGAGA	TCATCAGCAG	AGTTCACCAC
1090	1100	1110	1120	1130	1140
AGGCATCTGG	TGTCTCTTGT	TGGTTATTGC	ATCGCCGGTG	CCAAAAGATT	GCTTGTCTAT
1150	1160	1170	1180	1190	1200
GAGTTTGTTT	CTAACAACAA	TCTCGAGCTT	CACCTCCATG	GCGAGGGACG	GCCTACAATG
1210	1220	1230	1240	1250	1260
GAATGGAGCA	CCAGATTGAA	GATTGCTCTT	GGATCTGCTA	AAGGACTTTC	TTATCTTCAT
1270	1280	1290	1300	1310	1320
GAAGATTGCA	ATCCTAAAT	CATTACCGT	GATATCAAGG	CTTCAAACAT	ATTGATAGAT
1330	1340	1350	1360	1370	1380
TTCAAGTTTG	AAGCTAAGGT	TGCTGATTTT	GGTCTTGCTA	AGATTGCTTC	TGATACAAAC
1390	1400	1410	1420	1430	1440
ACGCATGTAT	CAACACGTGT	GATGGGAACC	TTTGGGTACT	TGGCTCCGGA	ATACGCTGCA
1450	1460	1470	1480	1490	1500
AGCGGAAAGC	TCACGGAGAA	GTCTGACGTT	TTCTCATTTG	GCGTTGTGCT	TTTGGAGCTC
1510	1520	1530	1540	1550	1560

FIG. 1D(1)

ALLEGRETTI, - 1931

Figure 1 C

ATTACTGGAC	GTCGACCCGT	TGATGCCAAC	AATGTCTATG	TAGATGACAG	CTTAGTTGAC
1570	1580	1590	1600	1610	1620
TGGGCACGAC	CATTGCTTAA	CCGAGCATCT	GAGCAAGGAG	ACTTTGAGGG	TTTAGCTGAT
1630	1640	1650	1660	1670	1680
GCAAAGATGA	ATAATGGGTA	TGACAGAGAG	GAGATGGCTC	GCATGGTTGC	TTGTGCTGCG
1690	1700	1710	1720	1730	1740
GCTTGTGTTT	GCCATTTCAGC	TCGCCGCGAGA	CCTCGCATGA	GCCAGATTGT	GCGTGCGTTA
1750	1760	1770	1780	1790	1800
GAAGGAAATG	TATCACTGTC	AGATCTTAAC	GAAGGGATGA	GACCAGGTCA	AAGCAATGTA
1810	1820	1830	1840	1850	1860
TACAGCTCAT	ACGGAGGAAG	CACCGATTAT	GAAGGGATGA	GACCAGGTCA	AAGCAATGTA
1870	1880	1890	1900	1910	1920
AAGTTTAGGA	AAATGGCACT	TGGAACCTCA	GAGTACAACG	CCACGGGTGA	GTACAGTAAT
1930	1940	1950	1960	1970	1980
CCGACCAGTG	ACTATGGACT	GTACCCGTCT	GGTTCAAGCA	GCGAGGGCCA	AACCACACGC
1990	2000	2010	2020	2030	2040
GAAATGGAGA	TGGGGAAGAT	TAAGAGAACC	GGTCAGGGTT	ATAGTGGACC	TTCTCTTTAA
2050	2060	2070	2080	2090	2100
ACCAGATGGG	AGAGAAATTG	AAGGGTGTTC	TTTCATTATT	TTTTTAAAC	TGTAAAGATA
2110	2120	2130	2140	2150	2160
TGAGAAAATT	GCCTTACTCT	AATTAAAACC	ACTACGATAT	AAGGTTATAA	TACGTTTTGA
2170	2180	2190	2200	2210	2220
ATTGGTTTTT	AAAAAAAAAA	AAAAAAAAAA

FIG. 1D(2)

LTLWSPCLLS	SPAASF*HSL	HLPF*FRSKE	ADMSSAPSPG	TGSPPSPPSN	STTTTPPPAS
APPPTTPSSP	PPPSTIPTSP	PPSSRSTPSA	PPSPPTPST	PGSPPLPQP	SPPAPTTPGS
PPAPVTPPTR	NPPPSVPGPP	SNPSREGGSP	RPPSSPSPPS	PSSDGLSTGV	VVGIAIGGVA
LLVIVTLICL	LCKKKRRRDE	EDAYYVPPPP	PPGPKAGGPY	GGQQQWRQQ	NATPPSDHVV
TSLPPPPKAP	SPPRQPPPPP	PPPFMSSSGG	SDYSDRPVLP	PPSPGLVLGF	SKSTFTYEEL
ARATNGFSEA	NLLQGGFGY	VHKGVLPSGK	EVAVKQLKVG	SGQGEREFQA	EVEIISRHH
RHLVSLVGYC	IAGAKRLLVY	EFVNNNNLEL	HLHGEGRPTM	EWSTRLKIAL	GSAGLSYLH
EDCNPKIIHR	DIKASNILID	FKFEAKVADF	GLAKIASDTN	THVSTRVMGT	FGYLAPEYAA
SGKLTEKSDV	FSFGVVLLEL	ITGRRPVDAN	NVYVDDSLVD	WARPLLNRAS	EQGDFEGLAD
AKMNNGYDRE	EMARMVACAA	ACVRHSARRR	PRMSQIVRAL	EGNVSLSDLN	EGMRPGQSNV
YSSYGSTDY	DSSQYNEDMK	KFRKMALGTQ	EYNATGEYSN	PTSDYGLYPS	GSSSEGQTTR
EMEMGKIKRT	GQGYSGPSL	TRWERN*RVF	FHYFFKTVKI	*ENCLTLIKT	TTI*GYNTF*
IGF*KKKKK					

FIG. 1E

```

ATTACTGGAC GTCGACCCGT TGATGCCAAC AATGTCTATG TAGATGACAG CTTAGTTGAC
1570 1580 1590 1600 1610 1620
TGGGCACGAC CATTGCTTAA CCGACCATCT GAGCAAGGAG ACTTTGAGGG TTTAGCTGAT
1630 1640 1650 1660 1670 1680
GCAAAGATGA ATAATGGGTA TGACAGAGAG GAGATGGCTC GCATGGTTGC TTGTGCTGCG
1690 1700 1710 1720 1730 1740
GCTTGTGTTT GCCATTCAGC TCGCCGCGAGA CCTCGCATGA GCCAGATTGT GCGTGCGTTA
1750 1760 1770 1780 1790 1800
GAAGGAAATG TATCACTGTC AGATCTTAAC GAAGGGATGA GACCAGGTCA AAGCAATGTA
1810 1820 1830 1840 1850 1860
TACAGCTCAT ACGGAGGAAG CACCGATTAT GACTCGAGCC AGTACAATGA AGACATGAAG
1870 1880 1890 1900 1910 1920
AAGTTTAGGA AAATGGCACT TGGAACTCAA GAGTACAACG CCACGGGTGA GTACAGTAAT
1930 1940 1950 1960 1970 1980
CCGACCAGTG ACTATGGACT GTACCCGTCT GGTTCAGCA GCGAGGGCCA AACCACACGC
1990 2000 2010 2020 2030 2040
GAAATGGAGA TGGGGAAGAT TAAGAGAACC GGTGAGGGTT ATAGTGGACC TTCTCTTTAA
2050 2060 2070 2080 2090 2100
ACCAGATGGG AGAGAAATTG AAGGGTGTTT TTTCATTATT TTTTAAAC TGTAAGATA
2110 2120 2130 2140 2150 2160
TGAGAAATT GCCTTACTCT AATTAAACC ACTACGATAT AAGGTTATAA TACGTTTGA
2170 2180 2190 2200 2210 2220
ATTGGTTTTT AAAAAAAAAA AAAAAAAAAA .....

```

Figure 1(d) (continued)

```

LTLWSPCLLS SPAASF*HSL HLPF*FRSKE ADMSSAPSPG TGSPPSPPSN STTTTPPPAS
APPTTTPSSP PPPSTIPTSP PPSSRSTPSA PPPSPPTPST PGSPPLPQP SPPAPTTPGS
PPAPVTEPTR NPPPSVPGFP SNPSREGGSP RPPSSPSPPS PSSDGLSTGV VVGIAIGGVA
LLVIVTLICL LCKKKRRRDE EDAYYVPPPP PPGPKAGGPY GGQQQQWRQQ NATPPSDHVV
TSLPPPPKAP SPPRQPPPPP PPFMSSSSGG SDYSDRPVLE PPSPGLVLGF SKSTFTYEEL
ARATNGFSEA NLLGQGGFGY VHKGVLPSGK EVAVKQLKVG SGQGEREFQA EVEIISRVHH
RHLVSLVGYC IAGAKRLLVY EFVENNNLEL HLHGEGRPTM EWSTRLKIAL GSAKGLSYLH
EDCNPKTIHR DIKASNILID FKFEAKVADF GLAKIASDTN THVSTRVMGT FGYLEPEYAA
SGKLTEKSDV FSFGVVLLEL ITGRRPVDAN NVYVDDSLVD WARPLNRRAS EQGDFEGLAD
AKMNNGYDRE EMARMVACAA ACVRHSARRR PRMSQIVRAL EGNVSLSDLN EGMRFQGSNV
YSSYGGSTDY DSSQYNEDMK KFRKMALGTQ EYNATGEYSN PTSYGLYPS GSSSEGQTR
EMEMGKIKRT GQYSGPSL* TRWERN*RVF FHYFFKTVKI *ENCLTIKTI TTI*GYNTF*
IGF*KKKKK

```

Figure 1(e)



PTO/PCT Rec'd 06 AUG 2002

7FS

July 29, 2002

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

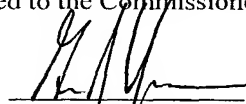
In re application of D. Goring and N. Silva
U.S. Application No. 10/069,304
Filed February 19, 2002
Proline-rich Extensin-like Receptor Kinases

Group No. NYA
Examiner NYA

(Atty. Docket No. P 25,762 USA)

CERTIFICATE OF MAILING

I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as first class mail in an envelope addressed to the Commissioner for Patents, Washington, D.C. 20231, on Monday, July 29, 2002.


Gene J. Yao

Commissioner for Patents
Washington, D.C. 20231

SUBMISSION OF SEQUENCE LISTING UNDER 37 C.F.R. 1.821

Sir:

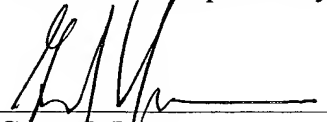
Attached herewith are a paper copy and an electronic copy of a Sequence Listing for the above-identified application. SEQ ID NOS.: 3, 4, and 5 are supported by Figures 1(c), 1(d), and 1(e), which were inadvertently not filed with International Application No. PCT/CA00/00966 (upon which the present application is a U.S. national stage thereof). These figures, however, were filed in U.S. provisional Application No. 60/159,122, which was incorporated in the present application by reference. A copy of the relevant pages of the '122 application is attached.

STATEMENT

In accordance with 37 CFR §1.821 relating to submissions containing Sequence Listings after the time of filing of the application, it is hereby certified that:

- (A) the contents of the computer-readable copy of the Sequence Listing submitted herewith and the paper copy of the Sequence Listing are the same;
- (B) all of the sequences are reflected in the application as filed (SEQ ID NOS.: 3, 4, and 5 are reflected in U.S. provisional Application No. 60/159,122, which was incorporated into the present application by reference); and
- (C) the present submission contains no new matter.

Submitted respectfully,


Gene J. Yao
Reg. No. 47,193
Attorney for Applicants

Synnestvedt & Lechner LLP
Suite 2600, Aramark Tower
1101 Market Street
Philadelphia, PA 19107-2950
(215) 923-4466 - Telephone
(215) 923-2189 - Facsimile

Proline-rich Extensin-like Receptor Kinases**FIELD OF THE INVENTION**

The invention relates to nucleic acid molecules and polypeptides involved in plant defense, and more particularly increasing plant resistance to wounding and pathogens.

5 BACKGROUND OF THE INVENTION

Receptor mediated signal perception and transduction in response to external stimuli are essential for growth and developmental processes of multicellular organisms (Mu et al., 1994). These extensively well characterized processes in animal systems involve receptor protein kinase molecules comprised of an extracellular signal perception domain, a
10 hydrophobic transmembrane domain attached to an intracellular domain that possesses kinase activity (Horn et al., 1994). In general, transmembrane signaling by receptor protein kinases requires binding of an appropriate ligand to the extracellular domain which induces receptor dimerization and alters the activity of the intracellular catalytic domain. This promotes phosphorylation of specific substrates thereby initiating a protein kinase signaling cascade
15 (Ullrich and Schlessinger, 1990). The majority of animal receptor protein kinases isolated to date contain tyrosine-specific kinase domains (Ullrich and Schlessinger, 1990), however, the transforming growth factor beta (TGF-beta) receptor (Lin et al., 1992) and the activin receptor (Dijke et al., 1993) possess kinase domains with serine/threonine phosphorylation activity.

Intracellular communication is also essential for the growth and development of
20 higher plants. The extensive knowledge of cell surface receptor signaling in animal systems has resulted in the isolation of several genes predicted to encode receptor-like protein kinases (RLKs). The characterized members of the RLK family share highly homologous catalytic domains with consensus sequences indicative of serine/threonine autophosphorylation activity, yet the extracellular domains of these receptors are very divergent (Braun and
25 Walker, 1996). Five different classes of plant receptor-like protein kinases have therefore been identified according to amino acid sequence similarity in the extracellular domains of these genes. The first class of receptor kinases, designated the S-domain class, have distinct extracellular domains homologous to the S-locus glycoprotein (SLG) (Nasrallah and Nasrallah, 1993). S-domain receptor kinases have several distinguishing features such as ten
30 conserved cysteine residues located proximal to the transmembrane domain in addition to other conserved residues implicated in the proper folding of the extracellular domain (Walker,

1994). Among this class of receptor kinases are the S-locus receptor kinases (SRKs) of Brassica expressed exclusively in reproductive tissues and implicated along with SLGs in controlling the sporophytic self-incompatibility response which normally inhibits self-pollination (Stein et al., 1991; Goring and Rothstein, 1992). Other receptor-like kinases of this type are represented in Arabidopsis by ARK1, ARK2, ARK3 (Tobias et al., 1992; Dwyer et al., 1994), in maize by ZmPK1 (Walker and Zhang, 1990) and by OsPK10 in rice (Zhao et al., 1994). The diversity in patterns of expression among members of the S-domain class shows that these plant receptor kinases are involved in mediating a variety of cellular signaling processes (Walker, 1994).

Another class of plant receptor kinases is the leucine-rich repeat (LRR) group which encodes proteins with extracellular domains containing 20-25 imperfect repeats of a 24 amino acid leucine-rich motif involved in peptide ligand recognition, cell adhesion and implicated in mediating protein-protein interactions (Braun and Walker, 1996; Wang et al., 1998). This class of plant receptor kinases include proteins such as CLAVATA1 (Clark et al., 1997) which is involved in regulating meristem and flower development in Arabidopsis, as well as proteins functioning in gamete development such as PRK1 of Petunia (Mu et al., 1994). The LRR class is represented in Arabidopsis by other receptor kinases such as ERECTA (Torii et al., 1996) which has been shown to be essential for proper plant and organ elongation, BRI1, a receptor involved in brassinosteroid signal transduction (Li and Chory, 1997), as well as TMK1 (Chang et al., 1992) and RLK5 (Walker, 1993) which may have more general roles in cellular signaling as suggested by their ubiquitous expression patterns in a variety of vegetative and reproductive tissues. Xa21, another member of the LRR class, has been implicated in pathogen recognition by providing resistance in rice to *Xanthomonas oryzae* pv. *oryzae* (Song et al., 1995).

The lectin-like class of plant receptor kinases is represented only in *Arabidopsis thaliana* by Ath.lecRK1 (Hervé et al., 1996) and LRK1 (Swarup et al., 1996). The extracellular domain of these receptor kinases share sequence similarity with lectins which are known carbohydrate binding proteins and implicated in the transduction of oligosaccharide signals in plant cellular communication processes (Hervé et al., 1996).

The two remaining classes of plant receptor kinases isolated in *Arabidopsis thaliana* include proteins with extracellular domains containing epidermal growth factor (EGF)-like motifs found in many proteins involved in extracellular interactions (WAK1; Kohorn et al.,

1992) as well as thaumatin-like domains homologous to PR5 proteins involved in pathogenesis (PR5K; Wang et al., 1996).

Plants remain very vulnerable to wounding and pathogens despite these advances. There is a need to identify other polypeptides, that help to protect plants. There is also a need for transgenic plants which overexpress these polypeptides and which have increased resistance to wounding and pathogens.

SUMMARY OF THE INVENTION

The invention relates to nucleic acid molecules and polypeptides involved in plant defense, and more particularly increasing plant resistance to wounding and pathogens.

We isolated a cDNA clone designated PERK1 (Proline-rich Extensin-like Receptor Kinase 1) which encodes a receptor kinase in *Brassica napus*. We define a new class of plant receptor kinases characterized by an extracellular domain rich in proline sharing sequence similarity to the extensin family of cell wall proteins. PERK1 is induced by both wounding and chemical elicitors which mimic a pathogen attack, showing a role for PERK1 in mediating a plant's defense response to mechanical and biological attack. Similar PERK nucleic acid molecules and polypeptides are found in other plants and cells.

The invention relates to an isolated nucleic acid molecule encoding a proline-rich, extensin-like receptor kinase (PERK) polypeptide, or a fragment of a PERK polypeptide having PERK activity. The molecule is a signaling molecule associated with the cell wall via its extensin-like extracellular domain and is involved in the transduction of extracellular stimuli into an intracellular response through a cytoplasmic kinase domain, thereby bridging the cell wall-plasma membrane continuum. The extracellular stimuli includes wounding or pathogen attack. The wounding can include a cut, a break, a tear, a fold or an insect wound. Typical pathogens include bacterial pathogens, fungal pathogens, *Sclerotinia sclerotiorum*, *Cylindrosporium conertricum*, *Phoma lingam*, *Pseudomonas syringae*, *Streptomyces scabies*, Blackleg, Whiterust, *Fusarium* Head Blight, Rust, Bunt, Leaf Spot, White mold, root rot or *Fusarium* ear rot

The invention also includes an isolated nucleic acid molecule encoding a PERK polypeptide, a fragment of a PERK polypeptide having PERK activity, or a polypeptide having PERK activity, comprising a nucleic acid molecule selected from the group consisting of:

(a) a nucleic acid molecule that hybridizes to a nucleic acid molecule consisting of [SEQ ID NO:1], or a complement thereof under low, moderate or high stringency hybridization conditions wherein the nucleic acid molecule encodes a PERK polypeptide or a polypeptide having PERK activity;

5 (b) a nucleic acid molecule degenerate with respect to (a), wherein the nucleic molecule encodes a PERK polypeptide or a polypeptide having PERK activity.

The hybridization conditions preferably include low stringency conditions of 1XSSC, 0.1% SDS at 50°C or high stringency conditions of 0.1XSSC, 0.1% SDS at 65°C. The invention also includes an isolated nucleic acid molecule encoding a PERK polypeptide, a
10 fragment of a PERK polypeptide having PERK activity, or a polypeptide having PERK activity, comprising a nucleic acid molecule selected from the group consisting of:

(a) the nucleic acid molecule of the coding strand shown in [SEQ ID NO:1], or a complement thereof;

15 (b) a nucleic acid molecule encoding the same amino acid sequence as a nucleotide sequence of (a); and

(c) a nucleic acid molecule having at least 17% identity with the nucleotide sequence of (a) and which encodes a PERK polypeptide or a polypeptide having PERK activity.

20 The invention also relates to an isolated nucleic acid molecule comprising a sequence that encodes a polypeptide having the sequence of SEQ ID NO:2, or the sequence of SEQ ID NO:2 with conservative amino acid substitutions.

The PERK polypeptide preferably consists of or comprises a PERK1 polypeptide.

The nucleic acid molecule preferably comprising all or part of a nucleotide sequence
25 shown in [SEQ ID NO:1 or 3] or a complement thereof. The nucleic acid may consist of the nucleotide sequence shown in [SEQ ID NO:1 or 3] or a complement thereof. The invention also includes a PERK1 nucleic acid molecule isolated from *Brassica*, or a fragment thereof. The *Brassica* may include *Brassica napus*, *Brassica juncea*, *Brassica rapa* or *Brassica oleracea*.

Another embodiment of the invention includes a recombinant nucleic acid molecule comprising a nucleic acid molecule of the invention and a constitutive promoter sequence or an inducible promoter sequence, operatively linked so that the promoter enhances transcription of the nucleic acid molecule in a host cell.

5 The nucleic acid molecule of the invention optionally includes genomic DNA, cDNA or RNA. The nucleic acid molecule is optionally chemically synthesized.

 The invention also includes an isolated nucleic acid molecule comprising a nucleic acid molecule selected from the group consisting of 8 to 10 nucleotides of the nucleic acid molecule of claim 6, 11 to 25 nucleotides of the nucleic acid molecule of claim 6 and 26 to 50
10 nucleotides of the nucleic acid molecule of claim 10. The nucleic acid molecule of the invention comprising at least 10, 15, 20, 30, 50 or 100 consecutive nucleotides of [SEQ ID NO:1 or 3] or a complement thereof. These sequences are useful as hybridization probes to detect PERK1 in a sample. They are also preferably useful as antisense oligonucleotides to inhibit gene expression. The invention also includes a nucleic acid molecule probe encoding
15 all or part (at least 10, 15, 20, 30, 50 or 100 amino acids) of PERK1 polypeptide.

 Another aspect of the invention includes a vector comprising a nucleic acid molecule of the invention. The vector optionally comprises a promoter selected from the group consisting of a super promoter, a 35S promoter of cauliflower mosaic virus, a chemical inducible promoter, a copper-inducible promoter, a steroid-inducible promoter and a tissue-specific promoter.
20

 Another variation of the invention includes a host cell comprising the recombinant nucleic acid molecule or the vector of the invention or progeny of the host cell. The host cell of the invention is optionally selected from the group consisting of a fungal cell, a yeast cell, a bacterial cell, a microorganism cell and a plant cell. The invention includes a method of
25 producing polypeptide, comprising culturing a host cell of the invention under conditions permitting expression of the polypeptide. The method preferably further includes isolating the protein from the cell or the cell medium.

 The invention also includes a plant, a plant part, a seed, a plant cell or progeny thereof comprising a recombinant nucleic acid molecule or vector of the invention. The plant part
30 optionally comprises all or part of a leaf, a flower, a stem, a root or a tuber. The plant, plant part, seed or plant cell is optionally selected from a species from the group consisting of

Brassica napus, *Brassica rapa*, *Brassica juncea*, *Brassica oleracea*, or from the family Brassicaceae, Arabidopsis, potato, tomato, tobacco, cotton, carrot, petunia, sunflower, strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley, sorghum and alfalfa. The plant may comprise a dicot plant or a monocot plant. The invention also includes an isolated polypeptide encoded by and/or produced from a nucleic acid molecule or the vector of the invention. The polypeptide is preferably an isolated PERK polypeptide or a fragment thereof having PERK activity. The invention also includes an isolated polypeptide, the amino acid sequence of which comprises at least ten consecutive residues of [SEQ ID NO:2]. The invention also includes an isolated immunogenic polypeptide, the amino acid sequence of which comprises at least ten consecutive residues of [SEQ ID NO:2]. The invention also includes an isolated polypeptide, the amino acid sequence of which comprises residues 1 to 137, 138 to 160 and 161 to 648 of [SEQ ID NO:2]. The polypeptide preferably comprises all or part of an amino acid sequence in [SEQ ID NO:2]. The invention also includes a polypeptide comprising or consisting of the sequence of SEQ ID NO:2 or 4, or the sequence of SEQ ID NO:2 or 4 with conservative amino acid substitutions. A variation of the invention comprises a polypeptide fragment of the PERK polypeptide, or a peptide mimetic of the PERK polypeptide (eg. PERK1). The polypeptide fragment of the invention preferably comprises at least 20 amino acids, which fragment has PERK activity. The fragment or peptide mimetic is preferably capable of being bound by an antibody to a polypeptide of the invention, such as PERK1. The polypeptide is optionally recombinantly produced. The invention includes an isolated and purified polypeptide comprising the amino acid sequence of a PERK polypeptide, wherein the polypeptide is encoded by a nucleic acid molecule that hybridizes under moderate or stringent conditions to a nucleic acid molecule in [SEQ ID NO:1 or 3], a degenerate form thereof or a complement. The invention also includes a polypeptide comprising an amino acid sequence having greater than 20% sequence identity to the polypeptide of [SEQ ID NO:1 or 3]. The polypeptide preferably comprises a PERK polypeptide. The polypeptide is preferably isolated from *Brassica*, for example, *Brassica napus* or *Brassica juncea* or *Brassica rapa* or *Brassica oleracea*.

The polypeptide preferably includes a kinase domain including at least 30% homology to the kinase domain of [SEQ ID NO.:2] and/or an extracellular domain including at least 20% homology to the extracellular domain of [SEQ ID NO.:2]. The invention includes an isolated nucleic acid molecule encoding a polypeptide the invention eg. [SEQ ID NO.:2].

The invention also includes an antibody that binds specifically to a polypeptide of the invention, in particular [SEQ ID NO.:2]. The antibody optionally comprises a monoclonal antibody or a polyclonal antibody. The invention also includes a purified polypeptide that binds specifically to an antibody that binds specifically to PERK1. Such proteins are useful,
5 for example, as a positive control in an assay utilizing the antibody.

The invention also includes an isolated nucleic acid molecule encoding a polypeptide that reduces the severity of wounding or pathogen attack in a plant, the polypeptide comprising:

- (a) an extracellular domain which recognizes an extracellular binding molecule whose
10 level is increased during the wounding or pathogen attack, the extracellular domain encoding a plurality of repeats selected from the group consisting of SPPPP, SPP, PP and PPP, wherein a plurality of the proline molecules are capable being glycosylated and/or hydroxylated;
- (b) a membrane domain operably connected to the extracellular domain, wherein the membrane domain is capable of extending across a cell membrane from the extracellular side
15 of the membrane to intracellular side of the membrane; and
- (c) a cytoplasmic domain operably connected to the membrane domain, wherein the cytoplasmic domain comprises a means for producing kinase activity when the extracellular binding molecule interacts with the extracellular domain.

The term "isolated nucleic acid" refers to a nucleic acid the structure of which is not
20 identical to that of any naturally occurring nucleic acid or to that of any fragment of a naturally occurring genomic nucleic acid spanning more than three separate genes. The term therefore covers, for example, (a) DNA which has the sequence of part of a naturally occurring genomic DNA molecules; (b) a nucleic acid incorporated into a vector or into the genomic DNA of a prokaryote or eukaryote, respectively, in a manner such that the resulting
25 molecule is not identical to any naturally occurring vector or genomic DNA; (c) a separate molecule such as cDNA, a genomic fragment, a fragment produced by reverse transcription of polyA RNA which can be amplified by PCR, or a restriction fragment; and (c) a recombinant nucleotide sequence that is part of a hybrid gene, i.e., a gene encoding a fusion protein. Specifically excluded from this definition are nucleic acids present in mixtures of (i)
30 DNA molecules, (ii) transfected cells, and (iii) cell clones, e.g., as these occur in a DNA library such as a cDNA or genomic DNA library.

BRIEF DESCRIPTION OF THE DRAWINGS

Preferred embodiments will be described in relation to the drawings in which:

Figure 1. Nucleotide and Deduced Amino Sequence. [SEQ ID NOS.: 1 & 2]

(A) Figure 1 shows the nucleic acid molecule of [SEQ ID NO.: 1] and the amino acid sequence of [SEQ ID NO.: 2].

In a preferred embodiment, the figure shows the nucleotide sequence corresponding to the PERK1 cDNA, with the predicted amino acid sequence presented as a single-letter code below the nucleotide sequence. Numbers to the left refer to nucleotide sequence and the 5' and 3' untranslated regions are presented in lower case letters. Potential N-glycosylation site (Asn-x-Ser/Thre) are indicated by dots above the Asn residues, and the extensin signature (Ser-Pro)₄ pentapeptide motif present in the extracellular domain is denoted in boldface type. The predicted membrane spanning region is marked by a solid underline. As defined by Hanks and Quinn (1991), the catalytic domain has been subdivided into 11 subdomains marked by dashed underlines and superscript roman numerals. The amino acids with a bracket underneath represent residues that are absolutely conserved, whereas amino acids with a wavy line underneath represent groups of conserved amino acids. The two regions marked by double underlines represent consensus sequences common amongst serine/threonine kinases.

(B) Structural features of the PERK1 polypeptide. A Kyte hydropathy plot (Kyte and Doolittle, 1982) of the amino acid sequence generated by DNAsis[®] software (Hitachi Software, San Bruno, CA) is shown, where increased hydrophobicity is denoted by positive values. The domains of the PERK1 protein are illustrated below. ECD, extracellular domain; TM, transmembrane domain.

(C) Shows the nucleic acid molecule of [SEQ ID NO.: 3] and the amino acid sequence of [SEQ ID NO.: 4]. These sequences are identical to nucleotides 1 to 1944 of [SEQ ID NO.: 1] and the corresponding amino acid sequence but they also include 5' and 3' untranslated nucleotide regions. (D) Shows the nucleic acid molecule of [SEQ ID NO.: 3] including the 5' and 3' untranslated regions. The start and stop codons are underlined and in light print. The nucleotide numbers correspond to the entire sequence and not only to the coding region. In a preferred embodiment, this is the coding sequence of PERK1.

(E) Figure 1 shows the amino acid sequence of [SEQ ID NO.: 5] including amino acids corresponding to the 5' and 3' untranslated regions. The starting methionine is in light print and the stop codon is indicated by an asterisk. It will be clear to those skilled in the art that one may only use the portion of the amino acid sequence between the start and stop codons.

5 **Figure 2. Genomic DNA Southern Blot Analysis of PERK1.**

Genomic DNA (5 micrograms) isolated from *Brassica napus* leaf tissue was digested with the indicated restriction enzymes, blotted and hybridized with a partial 1.5 kb PERK1 cDNA probe under varying conditions of stringency. DNA markers are indicated in kilobases.

(A) Genomic DNA gel blot analysis under low stringency conditions.

10 (B) Genomic DNA gel blot analysis under high stringency conditions.

Figure 3. Expression of PERK1 cDNA

(A) RNA gel blot analysis of PERK1 transcripts from total RNA extracted from various *Brassica napus* tissues. The blot was hybridized with the full length PERK1 cDNA probe and the expected ~2.2kb PERK1 transcript was detected (upper panel).

15 (B) The blot was subsequently probed with 18S rRNA as an internal control for even loading (lower panel).

Figure 4. Wound-Inducible Accumulation of PERK1 mRNA in *Brassica napus* Leaf and Stem Tissue.

(A) Fully expanded leaves were wounded by punching out discs around the perimeter of the leaf blade. Wounds mimic injury inflicted on plants in the field as a result of insect attack or other mechanical damage. Total RNA was extracted at various time intervals after treatment, subjected to Northern blot analysis and probed with full length PERK1 cDNA (bold-face arrow). The blot was reprobed with cyclophilin used as an internal control for even loading (open-face arrow). The graph represents the expression profile of PERK1 in response to wounding corrected against levels of cyclophilin expression. Error bars represent the standard error derived from two independent experiments. Control unwounded leaf tissue represented by 0 hr time point.

(B) Northern blot showing a time-course induction of PERK1 mRNA accumulation in wounded stem tissue. Total RNA harvested at the indicated time points was blotted and hybridized against the full length PERK1 coding sequence (bold-face arrow). The cyclophilin

loading control (open-face arrow) was used to normalize levels of PERK1 mRNA accumulation represented graphically. Error bars represent the standard error derived from two independent experiments. Control unwounded stem tissue represented by 0 hr time point.

Figure 4a. Wound-Inducible Accumulation of PERK1 mRNA in *Brassica napus* Leaf Disc Tissue.

The effects of a more localized wound defense response on the levels of PERK1 mRNA accumulation was investigated using wounded leaf discs. Total RNA was extracted from leaf discs at various time intervals, subjected the Northern blot analysis and probed with the full length PERK1 cDNA (bold-face arrow). The blot was reprobbed with cyclophilin as an internal control for even loading (open-face arrow). The graph represents the steady state levels of PERK1 mRNA in response to wounding corrected against levels of cyclophilin expression. Control unwounded leaf tissue is represented by the 0hr time point.

Figure 4b. Wound-Inducible Accumulation of PERK1 mRNA in *Brassica napus* Leaf and Stem Tissue.

(A) Fully expanded leaves were wounded by rubbing the undersides with abrasive sand paper. Total RNA was extracted from the leaf at various time intervals after treatment, subjected the Northern blot analysis and probed with the full length PERK1 cDNA (bold-face arrow). The blot was reprobbed with cyclophilin as an internal control for even loading (open-face arrow). The graph represents the steady state levels of PERK1 mRNA in response to wounding corrected against levels of cyclophilin expression. Control unwounded leaf tissue is represented by the 0hr time point.

(B) Northern blot analysis showing a time-course induction of PERK1 mRNA accumulation in stem wounded by rubbing with abrasive sand paper. Total RNA harvested at the indicated time points was blotted and hybridized against the full length PERK1 cDNA (bold-face arrow). The cyclophilin loading control (open-face arrow) was used to normalize levels of PERK1 mRNA accumulation represented graphically. Control unwounded stem tissue represented by 0hr time point.

Figure 4c. Wound-Inducible Accumulation of PERK1 mRNA in *Brassica napus* Root Tissue.

Root tissue from hydroponically grown *B. napus* plants was used to investigate whether levels of PERK1 mRNA increase in response to a wounding stimulus. Roots were wounded by

slicing tissue into 3cm sections and incubating on filter paper moistened with 20mM phosphate buffer supplemented with chloramphenicol. Total RNA extracted at various time intervals after treatment, was subjected the Northern blot analysis and probed with the full length PERK1 cDNA (bold-face triangle). The blot was reprobed with cyclophilin as an internal control for even loading (open-face triangle). The graph represents the expression profile of PERK1 in response to wounding corrected against levels of cyclophilin expression. Control unwounded root tissue is represented by 0hr time point

Figure 5. Effects of 50 micromolar Methyl Jasmonate (MeJA) on PERK1 mRNA Accumulation in Treated *Brassica napus* Leaf and Stem Tissue.

- 10 (A) *Brassica napus* plants were thoroughly sprayed with a 50 micromolar MeJA solution, and leaf tissue subsequently harvested at different time intervals after treatment. Total RNA prepared from treated leaf tissue was subjected to Northern blot analysis and probed with full length PERK1 cDNA (open-face triangle). Control plant (0 hr) was treated with the carrying solution minus the chemical inducer (0.1% [v/v] ethanol for MeJA). The blot was reprobed with cyclophilin used as an internal control for even loading (bold-face triangle). The graph represents a corrected profile for the levels of PERK1 mRNA accumulation in response to treatment with MeJA normalized against levels of cyclophilin expression.

- 15 (B) Northern blot showing a time-course induction of PERK1 mRNA accumulation in MeJA treated stem tissue. Total RNA harvested at the indicated time points was blotted and hybridized against the full length PERK1 coding sequence (open-face triangle). The cyclophilin loading control (bold-face triangle) was used to normalize levels of PERK1 mRNA accumulation represented graphically.

Figure 6. Effects of 4mM Salicylic Acid (SA) on PERK1 mRNA Accumulation in Treated *Brassica napus* Leaf and Stem Tissue.

- 25 (A) *Brassica napus* plants were thoroughly sprayed with a 4mM SA solution, and leaf tissue subsequently harvested at different time intervals after treatment. Total RNA prepared from treated leaf tissue was subjected to Northern blot analysis and probed with full length PERK1 cDNA (open-face triangle). Control plant (0 hr) was treated with the carrying solution minus the chemical inducer (5mM phosphate buffer, pH7). The blot was reprobed with cyclophilin used as an internal control for even loading (open-face triangle). The graph represents a
- 30

corrected profile for the levels of PERK1 mRNA accumulation in response to treatment with SA normalized against levels of cyclophilin expression.

(B) Northern blot showing a time-course induction of PERK1 mRNA accumulation in SA treated stem tissue. Total RNA harvested at the indicated time points was blotted and hybridized against the full length PERK1 coding sequence (open-face triangle). The cyclophilin loading control (bold-face triangle) was used to normalize levels of PERK1 mRNA accumulation represented graphically.

Figure 7. *B. napus* leaf tissue was used to investigate whether levels of PERK1 mRNA increase in response to treatment with a fungal pathogen *Sclerotinia sclerotiorum*. *Brassica napus* leaves were excised from the plant, placed in orchid tubes with water and incubated in closed aluminum trays at room temperature under fluorescent light to generate a humid environment. The leaves were inoculated with fungal agar plugs and incubated for the indicated time points, after which the tissue was harvested and immediately frozen in liquid nitrogen for further analysis. Control time points were inoculated with agar plugs only. Total RNA was extracted and Northern blot analysis was subsequently performed. Blots were probed with the full length PERK1 cDNA and then with cyclophilin as a control for even loading. Throughout the experiment, it was noted that the fungal agar plugs began to adhere well to the tissue 7hrs following inoculation and by 15hrs post-inoculation macroscopic lesions were apparent on the surface of the leaves. The corrected profile for PERK1 mRNA accumulation in response to this fungal pathogen treatment shows that there was a 2.5 fold induction in the levels of PERK1 mRNA 10hrs following treatment.

Figure 8. Proposed pathway mediating PERK1 expression in response to wounding, MeJA and SA treatments.

Figure 9. Analysis of Kinase Activity and Phosphoamino Acid Analysis of Recombinant

PERK1 Protein. (A) Purification of recombinant mutant (K-E) and wild type catalytic domain fusion proteins. Extracts of *E. coli* cells harboring the fusion proteins were purified on MBP amylose resin. Proteins were visualized by Coomassie blue staining. The mutant CD fusion protein is present in lane 1 and the wild type CD fusion protein is represented in lane 2.

(B) Western blot analysis of recombinant fusion proteins. Western blot analysis using an anti-MBP antibody was performed on purified Cdm_{ut}_{K-E} and CD_{wt} to confirm the identity of

these proteins. Lane 1 represents detection of the mutant fusion protein and the signal obtained in lane 2 verifies the identity of the wild type fusion protein.

(C) Autoradiogram of autophosphorylation assay of recombinant PERK1 CDwt and Cdmut_{K-E} fusion proteins. Detection of a signal in lane 1 proves that the wild type catalytic domain of PERK1 is capable of autophosphorylation. Absence of a signal for the Cdmut_{K-E} in lane 2 confirms that the result obtained for CDwt is specific to the kinase activity of PERK1 and not due to bacterial kinase contamination.

(D) Phosphoamino acid analysis of autophosphorylated PERK. γ ³²P-labeled CDwt was hydrolyzed with HCl and subjected to two-dimensional electrophoresis. Presence of radiolabeled phospho-serine (pS) and phospho-threonine (pT) confirms that PERK1 encodes an active protein kinase with serine/threonine substrate specificity.

Figure 10. Shows sequence identity of PERK1 to polypeptides from the Arabidopsis genome sequencing project.

Figure 11. Shows the nucleic acid molecule of [SEQ ID NO.: 6] and the amino acid sequence of [SEQ ID NO.: 7].

In a preferred embodiment, the figure shows the sequence of the predicted Arabidopsis gene - Accession number AAC98010

A) Genomic Sequence. The predicted open reading frame is underlined. The start codon (ATG) and stop codon (TGA) are double underlined.

B) Translation of the predicted open reading frame. The transmembrane domain is underlined.

Figure 12. Shows the nucleic acid molecule of [SEQ ID NO.: 8] and the amino acid sequence of [SEQ ID NO.: 9].

In a preferred embodiment, the figure shows the sequence of the predicted Arabidopsis gene - Accession number AAD15491

A) Genomic Sequence. The predicted open reading frame is underlined. The start codon (ATG) and stop codon (TGA) are double underlined.

B) Translation of the predicted open reading frame. The transmembrane domain is underlined.

Figure 13. Shows the nucleic acid molecule of [SEQ ID NO.: 10] and the amino acid sequence of [SEQ ID NO.: 11].

In a preferred embodiment, the figure shows the sequence of the predicted Arabidopsis gene - Accession number CAA18823.

A) Genomic Sequence. The predicted open reading frame is underlined. The start codon (ATG) and stop codon (TGA) are double underlined.

5 B) Translation of the predicted open reading frame. The transmembrane domain is underlined

Figure 14. Shows the nucleic acid molecule of [SEQ ID NO.: 12] and the amino acid sequence of [SEQ ID NO.: 13].

In a preferred embodiment, the figure shows the sequence of the predicted Arabidopsis gene - Accession number CAA18590

10

A) Genomic Sequence. The predicted open reading frame is underlined. The start codon (ATG) and stop codon (TGA) are double underlined.

B) Translation of the predicted open reading frame. The transmembrane domain is underlined.

15 DETAILED DESCRIPTION OF THE INVENTION

In this study, we report the isolation and preliminary characterization of PERK1 nucleic acid molecules and polypeptides, and in particular PERK1 cDNA which encodes a novel receptor-like protein kinase in *Brassica napus*. PERK polypeptides represent a novel class of receptor kinases in higher plants.

20 Protein kinases play important roles in plant defense (Zhou et al., 1995; Usami et al., 1995; Suzuki and Shinhi, 1995). Significant homology is shared between the extracellular domain of PERK1 and both extensin and proline rich proteins. PERK1 mediates plant responses to mechanical wounding (for example insect attack) and pathogen attack. PERK polypeptides preferably include a catalytic domain. PERK polypeptides are also preferably

25 signaling molecules associated with the cell wall via their extensin-like extracellular domain and involved in the transduction of extracellular stimuli (e.g. wounding, pathogen attack) into an intracellular response through a transmembrane domain and a cytoplasmic kinase domain, thereby bridging the cell wall - plasma membrane continuum.

In general, plants challenged by mechanical wounding or pathogen attack induce rapid

30 expression of genes (i.e. proteinase inhibitor (*pin*) and pathogenesis related (*PR*) genes

respectively) that are expressed locally as well as systemically in unaffected parts of the plant (Yang et al., 1997). Increased levels of extensin transcripts as a result of mechanical wounding have been well established in many other systems (Sauer et al., 1990; Shirsat et al., 1996). For example, in *Brassica napus* leaf and stem tissue, wound induction of PERK1 mRNA accumulation is a very rapid response detectable within 15 min following injury (Figure 4, 4a-b). Increased levels of PERK1 mRNA were also detected in wounded root tissue within 5 min following treatment (Figure 4c). MeJA (the methyl ester of the plant growth regulator jasmonic acid (JA)) is involved in the signal transduction pathway regulating gene activation upon wounding. Steady state levels of PERK1 mRNA remain unaffected by exogenously applied MeJA (Figure 5) which shows that the inducibility of PERK1 by wounding occurs via a MeJA-independent pathway (Figure 7). Studies conducted by Titarenko et al. (1997) addressing the role of JA in mediating wound responses support the existence of multiple distinct wound signal transduction pathways. Exogenously applied JA was able to induce only a subset of wound responsive genes in *Arabidopsis* which ultimately resulted in a stronger systemic accumulation in wounded plants. Conversely, a second set of wound responsive genes showing a stronger induction locally in wounded tissue showed no substantial accumulation upon JA treatment. In conjunction with the pattern of PERK1 mRNA accumulation in response to wounding and MeJA, it appears that plants respond to wounding by two distinct wound signal transduction pathways: one which does not require JA and is primarily responsible for gene activation in the vicinity of the wound site and the other which involves JA perception and activates gene expression both locally and systemically to the wound site (Titarenko et al., 1997).

Many of the inducible defense responses are not exclusive to mechanical wounding but are also initiated by pathogen attack. The similarity between responses to wounding and pathogen attack are not surprising since mechanical damage often precedes pathogen infection and conversely, mechanical damage may often result from a pathogen or insect attack (Truernit et al., 1996). Salicylic acid has been implicated in having an important role in the signal transduction pathway leading to systemic acquired resistance (SAR) (Penninckx et al., 1996). Steady state levels of PERK1 mRNA also accumulated in *B. napus* leaf and stem tissue upon exogenous application of 4mM SA (Figure 6). Collectively, the profiles of PERK1 mRNA accumulation in response to wounding, MeJA and SA are not entirely surprising. PERK1 induction is rapid in response to wounding (Figure 4, 4a-c) and the lack

of PERK1 transcript accumulation in response to MeJA (Figure 5) shows a pathway for wound mediated induction of PERK1 that is independent of MeJA (Figure 7). The pronounced and rapid induction of PERK1 in response to exogenous SA (Figure 6) supports other studies showing that SA is known to inhibit wound responsive genes that are regulated by a MeJA-dependent pathway (Peña-Cortés et al., 1993; Doares et al., 1995). Therefore, it is unlikely that both MeJA and SA would induce PERK1 mRNA accumulation given that these pathways are known to be antagonistic (Peña-Cortés et al., 1993). Nevertheless, the rapid induction of PERK1 during these treatments shows a role early on in a plant's defense signaling pathway.

10 Characterization of PERK1

Genomic Southern blot analysis under low and high stringency conditions revealed that PERK1 is a single copy gene in the *Brassica* genome (Figure 2). PERK1 is ubiquitously expressed at high levels in stem, petal and pistil tissue and is less abundant in root, leaf and anther tissues (upper panel).

15 The deduced amino acid sequence of PERK1 shows that it is a transmembrane receptor kinase with a distinct extracellular, transmembrane and cytoplasmic domain (Figure 1). The extracellular domain of PERK1 shows sequence similarity to plant cell wall proline-rich proteins and extensins which comprise a family of hydroxyproline-rich glycoproteins (HRGPs). Extensins are particularly abundant proteins in plant cell walls and are very rich in proline and serine as well as in combinations of valine, tyrosine, lysine and/or histidine residues. The distinctive characteristic of dicot extensins is their repetitive (Ser-Pro)₄ pentapeptide blocks. Although extensins are synthesized as soluble precursors, the majority of proline residues are hydroxylated and both the hydroxylated proline as well as the serine residues of these proteins are glycosylated by post-translational modifications (Cassab, 1998). 20 When secreted to the plant cell wall, extensins become rapidly insoluble, presumably due to the formation of covalent isodityrosine bridges (Cassab, 1998). Although extensins have been proposed to be structural cell wall proteins and important in development, they have also been directly implicated in plant defense against mechanical wounding (Shirsat et al., 1996) and pathogen attack (Corbin et al., 1987; Showalter, 1993). The catalytic domain of PERK1 25 possesses all of the invariant residues necessary for kinase activity and sequence similarity in subdomains VI and VIII to amino acid consensus sequences characteristic of serine/threonine kinases shows a role for PERK1 in plant signal transduction (Hanks and Quinn, 1991). Given 30

the similarity of PERK1 in the extracellular domain to the extensin family of cell wall proteins, PERK1 can detect changes to the cell wall through mechanical damage or pathogen attack and then pass the signal onto the cell. The cell can then respond to the attack with its defence mechanisms.

- 5 PERK nucleic acid molecules and proteins also have sequence identity and similarity to proline rich proteins as well. About 40% of PERK1's extracellular domain is comprised of proline. Below is a list of the Arabidopsis clones and their respective proline composition (Table below).

Table 1

10 % Prolines in Extracellular Domains

Gene	% Proline in ECD
PERK1	56/137 = 41%
CA18590	105/279 = 38%
AAC98010	85/246 = 34%
CAA18823	51/179 = 28%
AAD15491	36/149 = 24%

It is possible to use all or part of a proline rich domain from an extensin or a proline rich protein (or similar regions) to replace all or part of PERK1's extracellular domain.

- 15 In summary, PERK1 is a unique plant protein in *Brassica napus* involved in wound and pathogen response which physically links the cell wall and plasma membrane. PERK1 is involved in the general perception and subsequent transduction of a wound and/or pathogen stimulus, ultimately triggering a plant's defense mechanisms and conferring broad protection against such stimuli. Preliminary characterization showed that levels of PERK1 mRNA accumulate rapidly in response to wounding and SA. Further characterization of PERK1
- 20 induction with respect to changes in levels of phosphorylation provides additional evidence for the unequivocal role of PERK1 in plant defense. Furthermore, transgenic analysis of plants expressing altered levels of PERK1 confirms the involvement of PERK1 in wound and pathogen signaling. PERK1 polypeptides and nucleic acid molecules may be isolated from

the Brassicaceae including *Arabidopsis*, *Brassica napus*, *Brassica rapa*, *Brassica juncea*, *Brassica oleracea*, and other plants such as potato, tomato, tobacco, cotton, carrot, petunia, sunflower, strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley, sorghum and alfalfa.

5 PERK Nucleic Acid Molecules and Polypeptides

The invention relates to PERK nucleic acid molecules and polypeptides which increase wounding resistance and pathogen resistance in cells and plants. These polypeptides preferably include an extracellular domain, a transmembrane domain and a cytoplasmic domain. The cytoplasmic domain preferably includes a region with kinase activity. The
10 kinase activity is involved in cellular signaling. The PERK nucleic acid molecules which encode PERK polypeptides are particularly useful for producing transgenic plants which have increased wounding and pathogen resistance compared to a wild type plant.

It will also be apparent that there are polypeptide and nucleic acid molecules from other organisms, such as those listed previously, that are similar to PERK polypeptides and
15 nucleic acid molecules. PERK polypeptides are useful in increasing wounding and pathogen resistance in a cell, preferably a plant cell, because they include extensin-like and proline-rich domains (this refers to a plurality of domains including multiple proline residues, which are preferably similar to those found in extensins), such as SPPPP, SPP and PPP which are capable of being hydroxylated in response to wounding or pathogen attack. Once
20 hydroxylated, extensins become rapidly insoluble which strengthens the cell wall and in response to pathogen attack helps agglutinate or prevent the spread of the pathogen to neighbouring plant cells.

The PERK nucleic acid molecules and polypeptides, as well as their role in plants were not known before this invention. The ability of these compounds to increase wounding
25 and pathogen resistance of transgenic host cells (particularly plant cells) and transgenic plants compared to wild type cells and plants was unknown.

All nucleotides and polypeptides which are suitable for use in the methods of the invention, such as the preparation of transgenic host cells or transgenic plants, are included within the scope of the invention. Genomic clones or cDNA clones are preferred for
30 preparation of transgenic cells and plants.

In a preferred embodiment, the invention relates to a cDNA encoding PERK polypeptides from *Brassica napus*. The cDNA sequence and the corresponding amino acid sequence for PERK1 is presented in Figure 1. The invention also includes splice variants of the nucleic acid molecules as well as polypeptides produced from the molecules.

5 Characterization of Nucleic Acid Molecules and Polypeptides

In one variation, the invention includes DNA sequences (and the corresponding polypeptide) including at least one of the sequences shown in figure 1 in a nucleic acid molecule of preferably about: less than 1000 base pairs, less than 1250 base pairs, less than 1500 base pairs, less than 1750 base pairs, less than 2000 base pairs, less than 2250 base
10 pairs, less than 2500 base pairs, less than 2750 base pairs or less than 3000 base pairs.

The coding region of the PERK1 nucleic acid molecule is as follows:

Table 2

Nucleic Acid Molecule	Start Nucleotide [brackets show corresponding amino acid nos.]	End Nucleotide [brackets show corresponding amino acid nos.]
PERK1 (coding region only)	1 (1)	1944 (648)
PERK1 Extracellular Domain	1 (1)	411 (137)
PERK1 Transmembrane Domain	412 (138)	480 (160)
PERK1 Cytoplasmic Domain	481 (161)	1944 (648)
PERK1 Kinase region	Same as cytoplasmic domain	Same as cytoplasmic domain

It will be apparent that these may be varied, for example, by shortening the 5'
15 untranslated region or shortening the nucleic acid molecule so that the 3' end nucleotide is in
a different position.

The discussion of the nucleic acid molecules, sequence identity, hybridization and other aspects of nucleic acid molecules included within the scope of the invention is intended to be applicable to either the entire nucleic acid molecule in figure 1 or the coding region of this molecule, shown in Table 2. One may use the entire molecule in figure 1 or only the coding region. Other possible modifications to the sequence are apparent.

Southern Blot Analysis shows that PERK1 is present as a single copy gene in *Brassica*. A Northern blot showed that PERK1 polypeptide was expressed in all tissues examined (root, stem, leaf, petal, anther and pistil). It is highly expressed in the stem, petal and pistil tissues and to a lower extent in the root, leaf and anther tissues.

10 The PERK1 Nucleic Acid Molecule and Polypeptide are Conserved in Plants

Sequence Identity

This is the first isolation of a nucleic acid molecule encoding a PERK polypeptide from plant species. Nucleic acid sequences having sequence identity to the PERK1 sequence are found in other species of *Brassica* such as *Brassica rapa*, *Brassica juncea*, and *Brassica oleracea* as well as other plants such as *Arabidopsis*, potato, tomato, tobacco, cotton, carrot, petunia, sunflower, strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley, sorghum and alfalfa. Sequences from *Brassica napus* and other plants are collectively referred to as "PERK" nucleic acid sequences and polypeptides. We isolate PERK nucleic acid molecules from plants having nucleic acid molecules that are similar to those in *Brassica napus*, such as beet, tomato, rice, cucumber, radish and other plants including *Arabidopsis*, potato, tobacco, cotton, carrot, petunia, sunflower, strawberries, spinach, lettuce, soybean, corn, wheat, rye, barley, sorghum and alfalfa. and using techniques described in this application. The invention includes methods of isolating these nucleic acid molecules and polypeptides as well as methods of using these nucleic acid molecules and polypeptides according to the methods described in this application, for example those used with respect to PERK1.

Table 3 below shows several sequences with sequence identity and sequence similarity to the PERK1 polypeptide. Where polypeptides are shown, a suitable corresponding DNA encoding the polypeptide will be apparent. These sequences code for polypeptides similar to portions of PERK1 polypeptide. The sequences in Table 3 are useful to make probes to identify full length sequences or fragments (from the listed species or other

species). They are useful to screen for functionally related cDNAs and genes. They are also useful to screen other tissues to see if they include all or part of the shown EST or similar sequences. The invention also relates to nucleic acid molecules including these EST sequences. In particular, the invention includes an isolated nucleic acid comprising a

5 sequence that hybridizes under moderate or stringent conditions to SEQ ID NOS: 1 or 3 or the complements thereof. The invention also includes an isolated nucleic acid molecule comprising a sequence having at least about: >60%, >70%, >80% or >90%, more preferably at least about: >95%, or >99% sequence identity to a sequence in Table 3. One skilled in the art would be able to design a probe based on a polypeptide or peptide fragment. The

10 invention includes nucleic acid molecules of about: 10 to 50 nucleotides, 50 to 200 nucleotides, 200 to 500 nucleotides, 500 to 1000 nucleotides, 1000 to 1500 nucleotides, 1500 to 1700 nucleotides, 1700 to 2000 nucleotides, 2000 to 2500 nucleotides or at least 2500 nucleotides and which include all or part of the sequences (or corresponding nucleic acid molecule) in Table 3. The invention also includes a nucleic acid molecule including the

15 sequences in Table 3 which encodes peptides and polypeptides of about: 10 to 50 amino acids, 50 to 200 amino acids, 200 to 500 amino acids, 500 to 750 amino acids or at least 750 amino acids. Possible modifications to these sequences will also be apparent. The polypeptide and nucleic acid molecules are also useful in research experiments or in bioinformatics to locate other sequences. The nucleic acid molecules and polypeptides

20 preferably provide PERK activity.

Table 3

Organism	Accession No.
<i>Arabidopsis thaliana</i>	AAC 98010 (Figure 11)
<i>Arabidopsis thaliana</i>	AAD 15491 (Figure 12)
<i>Arabidopsis thaliana</i>	CAA 18823 (Figure 13)
<i>Arabidopsis thaliana</i>	CAA 18590 (Figure 14)

The regions of importance include the extracellular domain (ECD), transmembrane domain

25 (TMD), and the catalytic domain. Amino acid positions are as follows:

AAC98010 – ECD: 1-247; TMD: 248-267; CD: 268 -732

AAD15491- ECD:1- 149 ; TMD – 150-171; CD - 172 -634

CAA18823 – ECD:1- 179 ; TMD – 180-194; CD - 195 -675

CAA18590 - ECD:1- 279 ; TMD – 280-302; CD - 303 -732

The invention includes the nucleic acid molecules from other plants as well as
5 methods of obtaining the nucleic acid molecules by, for example, screening a cDNA library
or other DNA collections with a probe of the invention (such as a probe comprising at least
about: 10 or preferably at least 15 or 30 nucleotides of PERK1 and detecting the presence of a
PERK nucleic acid molecule. Another method involves comparing the PERK1 sequences to
other sequences, for example using bioinformatics techniques such as database searches or
10 alignment strategies, and detecting the presence of a PERK nucleic acid molecule or
polypeptide. The invention includes the nucleic acid molecule and/or polypeptide obtained
according to the methods of the invention. The invention also includes methods of using the
nucleic acid molecules, for example to make probes, in research experiments or to transform
host cells or make transgenic plants. These methods are as described below.

The polypeptides encoded by the PERK nucleic acid molecules in other species will have amino acid sequence identity to the PERK1 sequence. Sequence identity may be at least about: >20%, >25%, >28%, >30%, >35%, >40%, >50% to an amino acid sequence shown in figure 1 (or a partial sequence thereof). Some polypeptides may have a sequence identity of at least about: >60%, >70%, >80% or >90%, more preferably at least about: >95%, >99% or >99.5% to an amino acid sequence in figure 1 (or a partial sequence thereof). Identity is calculated according to methods known in the art. Sequence identity (nucleic acid and protein) is most preferably assessed by the algorithm of the Fasta 3 program, using the following default parameter settings: gap penalty (open) = -12 (protein) -16 (DNA), gap penalty (extension) = -2 (protein) -4 (DNA), protein weight matrix = BLOSUM 62. (The reference for FASTA 3 is W. R. Pearson and D. J. Lipman (1988), "Improved Tools for Biological Sequence Analysis", PNAS 85:2444- 2448, and W. R. Pearson (1990) "Rapid and Sensitive Sequence Comparison with FASTP and FASTA" Methods in Enzymology 183:63-98). The invention also includes modified polypeptide from plants which have sequence identity at least about: >20%, >25%, >28%, >30%, >35%, >40%, >50%, >60%, >70%, >80% or >90% more preferably at least about >95%, >99% or >99.5%, to the PERK sequence in figure 1 (or a partial sequence thereof). Modified polypeptide molecules are discussed below. Preferably about: 1, 2, 3, 4, 5, 6 to 10, 10 to 25, 26 to 50 or 51 to 100, or 101 to 250 nucleotides or amino acids are modified.

Nucleic Acid Molecules and Polypeptides Similar to PERK1

Those skilled in the art will recognize that the nucleic acid molecule sequences in figure 1 are not the only sequences which may be used to provide increased PERK activity in plants. The genetic code is degenerate so other nucleic acid molecules which encode a polypeptide identical to an amino acid sequence in figure 1 may also be used. The sequence of the other nucleic acid molecules of this invention may also be varied without changing the polypeptide encoded by the sequence. Consequently, the nucleic acid molecule constructs described below and in the accompanying examples for the preferred nucleic acid molecules, vectors, and transformants of the invention are merely illustrative and are not intended to limit the scope of the invention.

The sequences of the invention can be prepared according to numerous techniques. The invention is not limited to any particular preparation means. For example, the nucleic acid molecules of the invention can be produced by cDNA cloning, genomic cloning, cDNA

synthesis, polymerase chain reaction (PCR), or a combination of these approaches (Current Protocols in Molecular Biology (F. M. Ausbel et al., 1989).). Sequences may be synthesized using well known methods and equipment, such as automated synthesizers. Nucleic acid molecules may be amplified by the polymerase chain reaction. Polypeptides may, for example, be synthesized or produced recombinantly.

Sequence Identity

The invention includes modified nucleic acid molecules with a sequence identity at least about: >17%, >20%, >30%, >40%, >50%, >60%, >70%, >80% or >90% more preferably at least about >95%, >99% or >99.5%, to a DNA sequence in figure 1 (or a partial sequence thereof). Preferably about 1, 2, 3, 4, 5, 6 to 10, 10 to 25, 26 to 50 or 51 to 100, or 101 to 250 nucleotides or amino acids are modified. Identity is calculated according to methods known in the art. Sequence identity is most preferably assessed by the algorithm of the FASTA 3 program. For example, if a nucleotide sequence (called "Sequence A") has 90% identity to a portion of the nucleotide sequence in Figure 1, then Sequence A will be identical to the referenced portion of the nucleotide sequence in Figure 1, except that Sequence A may include up to 10 point mutations, such as substitutions with other nucleotides, per each 100 nucleotide of the referenced portion of the nucleotide sequence in Figure 1. Nucleotide sequences functionally equivalent to the PERK1 sequence can occur in a variety of forms as described below. Polypeptides having sequence identity may be similarly identified.

The polypeptides encoded by the homologous PERK nucleic acid molecule in other species will have amino acid sequence identity at least about: >20%, >25%, >28%, >30%, >40% or >50% to an amino acid sequence shown in figure 1 (or a partial sequence thereof). Some plant species may have polypeptides with a sequence identity of at least about: >60%, >70%, >80% or >90%, more preferably at least about: >95%, >99% or >99.5% to all or part of an amino acid sequence in figure 1 (or a partial sequence thereof). Identity is calculated according to methods known in the art. Sequence identity is most preferably assessed by the FASTA 3 program. Preferably about: 1, 2, 3, 4, 5, 6 to 10, 10 to 25, 26 to 50 or 51 to 100, or 101 to 250 nucleotides or amino acids are modified.

The invention includes nucleic acid molecules with mutations that cause an amino acid change in a portion of the polypeptide not involved in providing PERK activity or an amino acid change in a portion of the polypeptide involved in providing PERK activity so that the mutation increases or decreases the activity of the polypeptide.

5 Hybridization

Other functional equivalent forms of the PERK nucleic acid molecules encoding nucleic acids can be isolated using conventional DNA-DNA or DNA-RNA hybridization techniques. These nucleic acid molecules and the PERK sequences can be modified without significantly affecting their activity.

10 The present invention also includes nucleic acid molecules that hybridize to one or more of the sequences in figure 1 (or a partial sequence thereof) or their complementary sequences, and that encode peptides or polypeptides exhibiting substantially equivalent activity as that of an PERK polypeptide produced by the DNA in figure 1. Such nucleic acid molecules preferably hybridize to all or a portion of PERK or its complement or all or a
15 portion of an EST of Table 3 under low, moderate (intermediate), or high stringency conditions as defined herein (see Sambrook et al. (Most recent edition) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.; Ausubel et al. (eds.), 1995, Current Protocols in Molecular Biology, (John Wiley & Sons, NY)). The portion of the hybridizing nucleic acids is typically at least 15 (e.g. 20, 25, 30 or
20 50) nucleotides in length. The hybridizing portion of the hybridizing nucleic acid is at least 80% e.g. at least 95% or at least 98% identical to the sequence or a portion or all of a nucleic acid encoding a PERK polypeptide, or its complement. Hybridizing nucleic acids of the type described herein can be used, for example, as a cloning probe, a primer (e.g. a PCR primer) or a diagnostic probe. Hybridization of the oligonucleotide probe to a nucleic acid sample
25 typically is performed under stringent conditions. Nucleic acid duplex or hybrid stability is expressed as the melting temperature or T_m , which is the temperature at which a probe dissociates from a target DNA. This melting temperature is used to define the required stringency conditions. If sequences are to be identified that are related and substantially identical to the probe, rather than identical, then it is useful to first establish the lowest
30 temperature at which only homologous hybridization occurs with a particular concentration of salt (e.g. SSC or SSPE). Then, assuming that 1% mismatching results in a 1 degree Celsius decrease in the T_m , the temperature of the final wash in the hybridization reaction is reduced

accordingly (for example, if sequences having greater than 95% identity with the probe are sought, the final wash temperature is decreased by 5 degrees Celsius). In practice, the change in T_m can be between 0.5 degrees Celsius and 1.5 degrees Celsius per 1% mismatch. Low stringency conditions involve hybridizing at about: 2XSSC, 0.1% SDS at 50°C. High stringency conditions are: 0.1XSSC, 0.1% SDS at 65°C. Moderate stringency is about 1X SSC 0.1% SDS at 60 degrees Celsius. The parameters of salt concentration and temperature can be varied to achieve the optimal level of identity between the probe and the target nucleic acid.

The present invention also includes nucleic acid molecules from any source, whether modified or not, that hybridize to genomic DNA, cDNA, or synthetic DNA molecules that encode the amino acid sequence of a PERK polypeptide, or genetically degenerate forms, under salt and temperature conditions equivalent to those described in this application, and that code for a peptide, or polypeptide that has PERK activity. Preferably the polypeptide has the same or similar activity as that of a PERK1 polypeptide. A nucleic acid molecule described above is considered to be functionally equivalent to a PERK nucleic acid molecule (and thereby having PERK activity) of the present invention if the polypeptide produced by the nucleic acid molecule displays the following characteristics: The defining feature of PERK polypeptides is the presence of a proline-rich domain, followed by a transmembrane domain, followed by a kinase domain. When tested, the kinase domain has serine/threonine kinase activity.

The invention also includes nucleic acid molecules and polypeptides having sequence similarity taking into account conservative amino acid substitutions. Sequence similarity (and preferred percentages) are discussed below.

Modifications to Nucleic Acid Molecule or Polypeptide Sequence

Changes in the nucleotide sequence which result in production of a chemically equivalent or chemically similar amino acid sequences are included within the scope of the invention. Variants of the polypeptides of the invention may occur naturally, for example, by mutation, or may be made, for example, with polypeptide engineering techniques such as site
5 directed mutagenesis, which are well known in the art for substitution of amino acids. For example, a hydrophobic residue, such as glycine can be substituted for another hydrophobic residue such as alanine. An alanine residue may be substituted with a more hydrophobic residue such as leucine, valine or isoleucine. A negatively charged amino acid such as aspartic acid may be substituted for glutamic acid. A positively charged amino acid such as
10 lysine may be substituted for another positively charged amino acid such as arginine.

Therefore, the invention includes polypeptides having conservative changes or substitutions in amino acid sequences. Conservative substitutions insert one or more amino acids which have similar chemical properties as the replaced amino acids. The invention includes sequences where conservative substitutions are made that do not destroy PERK
15 activity. The preferred percentage of sequence similarity for sequences of the invention includes sequences having at least about: 50% similarity to PERK1. The similarity may also be at least about: 60% similarity, 75% similarity, 80% similarity, 90% similarity, 95% similarity, 97% similarity, 98% similarity, 99% similarity, or more preferably at least about 99.5% similarity, wherein the polypeptide has PERK activity. The invention also includes
20 nucleic acid molecules encoding polypeptides, with the polypeptides having at least about: 50% similarity to PERK1. The similarity may also be at least about: 60% similarity, 75% similarity, 80% similarity, 90% similarity, 95% similarity, 97% similarity, 98% similarity, 99% similarity, or more preferably at least about 99.5% similarity, wherein the polypeptide has PERK activity, to an amino acid sequence in figure 1 (or a partial sequence thereof)
25 considering conservative amino acid changes, wherein the polypeptide has PERK activity. Sequence similarity is preferably calculated number of similar amino acids in a multiple alignment expressed as a percentage of the shorter of the two sequences in the alignment. The multiple alignment is preferably constructed using the algorithm of the FASTA 3 program, using the following parameter settings: gap penalty (open) = -12(protein) -16 (DNA), gap
30 penalty (extension) = -2 (protein) -4 (DNA) , protein weight matrix = BLOSUM 62. (The reference for FASTA 3 is W. R. Pearson and D. J. Lipman (1988), "Improved Tools for Biological Sequence Analysis", PNAS 85:2444- 2448, and W. R. Pearson (1990) "Rapid and

Sensitive Sequence Comparison with FASTP and FASTA" Methods in Enzymology 183:63-98).

Polypeptides comprising one or more d-amino acids are contemplated within the invention. Also contemplated are polypeptides where one or more amino acids are acetylated at the N-terminus. Those of skill in the art recognize that a variety of techniques are available for constructing polypeptide mimetics with the same or similar desired PERK activity as the corresponding polypeptide compound of the invention but with more favorable activity than the polypeptide with respect to solubility, stability, and/or susceptibility to hydrolysis and proteolysis. See, for example, Morgan and Gainor, Ann. Rep. Med. Chem., 24:243-252 (1989). Examples of polypeptide mimetics are described in U.S. Patent Nos. 5,643,873. Other patents describing how to make and use mimetics include, for example in, 5,786,322, 5,767,075, 5,763,571, 5,753,226, 5,683,983, 5,677,280, 5,672,584, 5,668,110, 5,654,276, 5,643,873. Mimetics of the polypeptides of the invention may also be made according to other techniques known in the art. For example, by treating a polypeptide of the invention with an agent that chemically alters a side group by converting a hydrogen group to another group such as a hydroxy or amino group. Mimetics preferably include sequences that are either entirely made of amino acids or sequences that are hybrids including amino acids and modified amino acids or other organic molecules.

The invention also includes hybrid nucleic acid molecules and polypeptides, for example where a nucleotide sequence from one species of plant is combined with a nucleotide sequence from another sequence of plant, mammal, bacteria or yeast to produce a fusion polypeptide. The invention includes a fusion protein having at least two components, wherein a first component of the fusion protein comprises a polypeptide of the invention, preferably a full length PERK polypeptide. The second component of the fusion protein preferably comprises a tag, for example GST, an epitope tag or an enzyme. The fusion protein may comprise lacZ.

The invention also includes polypeptide fragments of the polypeptides of the invention which may be used to confer PERK activity if the fragments retain activity. The invention also includes polypeptides fragments of the polypeptides of the invention which may be used as a research tool to characterize the polypeptide or its activity. Such polypeptides preferably consist of at least 5 amino acids. In preferred embodiments, they may consist of 6 to 10, 11 to 15, 16 to 25, 26 to 50, 51 to 75, 76 to 100 or 101 to 250 amino acids

of the polypeptides of the invention (or longer amino acid sequences). The fragments preferably have PERK activity. Fragments may include sequences with one or more amino acids removed, for example, C-terminus amino acids in a PERK sequence.

5 The invention also includes a composition comprising all or part of an isolated PERK nucleic acid molecule (preferably PERK1) of the invention and a carrier, preferably in a composition for plant transformation. The invention also includes a composition comprising an isolated PERK polypeptide (preferably PERK1) and a carrier, preferably for studying polypeptide activity.

Recombinant Nucleic Acid Molecules

10 The invention also includes recombinant nucleic acid molecules preferably a PERK1 sequence of figure 1 comprising a nucleic acid molecule of the invention and a promoter sequence, operatively linked so that the promoter enhances transcription of the nucleic acid molecule in a host cell (the nucleic acid molecules of the invention may be used in an isolated native gene or a chimeric gene, for example, where a nucleic acid molecule coding region is
15 connected to one or more heterologous sequences to form a gene. The promoter sequence is preferably a constitutive promoter sequence or an inducible promoter sequence, operatively linked so that the promoter enhances transcription of the DNA molecule in a host cell. The promoter may be of a type not naturally associated with the cell such as a super promoter, a 35S cauliflower mosaic virus promoter, a chemical inducible promoter, a copper-inducible
20 promoter, a steroid-inducible promoter and a tissue specific promoter .

A recombinant nucleic acid molecule for conferring PERK activity may also contain suitable transcriptional or translational regulatory elements. Suitable regulatory elements may be derived from a variety of sources, and they may be readily selected by one with ordinary skill in the art. Examples of regulatory elements include: a transcriptional promoter and enhancer or RNA polymerase binding sequence, a ribosomal binding sequence, including a translation initiation signal. Additionally, depending on the vector employed, other genetic elements, such as selectable markers, may be incorporated into the recombinant molecule. Markers facilitate the selection of a transformed host cell. Such markers include genes associated with temperature sensitivity, drug resistance, or enzymes associated with phenotypic characteristics of the host organisms.

Nucleic acid molecule expression levels are controlled with a transcription initiation region that regulates transcription of the nucleic acid molecule or nucleic acid molecule fragment of interest in a plant, bacteria or yeast cell. The transcription initiation region may be part of the construct or the expression vector. The transcription initiation domain or promoter includes an RNA polymerase binding site and an mRNA initiation site. Other regulatory regions that may be used include an enhancer domain and a termination region. The regulatory elements described above may be from animal, plant, yeast, bacteria, fungus, virus or other sources, including synthetically produced elements and mutated elements.

Methods of modifying DNA and polypeptides, preparing recombinant nucleic acid molecules and vectors, transformation of cells, expression of polypeptides are known in the art. For guidance, one may consult the following US patent nos. 5,840,537, 5,850,025, 5,858,719, 5,710,018, 5,792,851, 5,851,788, 5,759,788, 5,840,530, 5,789,202, 5,871,983, 5,821,096, 5,876,991, 5,422,108, 5,612,191, 5,804,693, 5,847,258, 5,880,328, 5,767,369, 5,756,684, 5,750,652, 5,824,864, 5,763,211, 5,767,375, 5,750,848, 5,859,337, 5,563,246, 5,346,815, and WO9713843. Many of these patents also provide guidance with respect to experimental assays, probes and antibodies, methods, transformation of host cells and regeneration of plants, which are described below. These patents, like all other patents, publications (such as articles and database publications) in this application, are incorporated by reference in their entirety.

Host Cells Including a PERK Nucleic Acid Molecule

In a preferred embodiment of the invention, a plant or yeast cell is transformed with a nucleic acid molecule of the invention or a fragment of a nucleic acid molecule and inserted in a vector.

5 Another embodiment of the invention relates to a method of transforming a host cell with a nucleic acid molecule of the invention or a fragment of a nucleic acid molecule, inserted in a vector. The invention also includes a vector comprising a nucleic acid molecule of the invention. The nucleic acid molecules can be cloned into a variety of vectors by means that are well known in the art. The recombinant nucleic acid molecule may be inserted at a
10 site in the vector created by restriction enzymes. A number of suitable vectors may be used, including cosmids, plasmids, bacteriophage, baculoviruses and viruses. Suitable vectors are capable of reproducing themselves and transforming a host cell. The invention also relates to a method of expressing polypeptides in the host cells. A nucleic acid molecule of the invention may be used to transform virtually any type of plant, including both monocots and
15 dicots. The expression host may be any cell capable of expressing PERK, such as a cell selected from the group consisting of a seed (where appropriate), plant cell, bacterium, yeast, fungus, protozoa, algae, animal and animal cell.

Levels of nucleic acid molecule expression may be controlled with nucleic acid molecules or nucleic acid molecule fragments that code for anti-sense RNA inserted in the
20 vectors described above.

Agrobacterium tumefaciens-mediated transformation, particle-bombardment-mediated transformation, direct uptake, microinjection, coprecipitation and electroporation-mediated nucleic acid molecule transfer are useful to transfer a PERK nucleic acid molecule into seeds (where appropriate) or host cells, preferably plant cells, depending upon the plant species.
25 The invention also includes a method for constructing a host cell capable of expressing a nucleic acid molecule of the invention, the method comprising introducing into said host cell a vector of the invention. The genome of the host cell may or may not also include a functional PERK gene. The invention also includes a method for expressing a PERK polypeptide such as a PERK1 in figure 1 in the host cell or a plant, plant part, seed or plant
30 cell of the invention, the method comprising culturing the host cell under conditions suitable for gene expression. The method preferably also includes recovering the expressed polypeptide from the culture.

The invention includes the host cell comprising the recombinant nucleic acid molecule and vector as well as progeny of the cell. Preferred host cells are fungal cells, yeast cells, bacterial cells, mammalian cells, bird cells, reptile cells, amphibious cells, microorganism cells and plant cells. Host cells may be cultured in conventional nutrient media. The media may be modified as appropriate for inducing promoters, amplifying genes or selecting transformants. The culture conditions, such as temperature, composition and pH will be apparent. After transformation, transformants may be identified on the basis of a selectable phenotype. A selectable phenotype can be conferred by a selectable marker in the vector.

10 Transgenic Plants and Seeds

Plant cells are useful to produce tissue cultures, seeds or whole plants. The invention includes a plant, plant part, seed, or progeny thereof including a host cell transformed with a PERK nucleic acid molecule such as a molecule in figure 1. The plant part is preferably a leaf, a stem, a flower, a root, a seed or a tuber.

15 The invention includes a transformed (transgenic) plant having increased PERK activity, the transformed plant containing a nucleic acid molecule sequence encoding for polypeptide activity and the nucleic acid molecule sequence having been introduced into the plant by transformation under conditions whereby the transformed plant expresses a PERK polypeptide in an active form.

20 The methods and reagents for producing mature plants from cells are known in the art. The invention includes a method of producing a genetically transformed plant which expresses PERK polypeptide such as a polypeptide in figure 1 by regenerating a genetically transformed plant from the plant cell, seed or plant part of the invention. The invention also includes the transgenic plant produced according to the method. Alternatively, a plant may be transformed with a vector of the invention.

25 The invention also includes a method of preparing a plant with increased PERK activity, the method comprising transforming the plant with a nucleic acid molecule which encodes a polypeptide of figure 1 or a polypeptide encoding a PERK polypeptide capable of increasing PERK activity in a cell, and recovering the transformed plant with increased PERK activity. The invention also includes a method of preparing a plant with increased PERK activity, the method comprising transforming a plant cell with a nucleic acid molecule such as

a molecule of figure 1 which encodes a PERK polypeptide capable of increasing PERK activity in a cell.

Overexpression of PERK leads to an improved ability of the transgenic plants to resist wounding or pathogen damage.

5 The plants whose cells may be transformed with a nucleic acid molecule of this invention and used to produce transgenic plants include, but are not limited to the

Target plants: *Brassica napus*, *Brassica rapa*, *Brassica juncea*, *Brassica oleracea*, or from the family Brassicaceae, Arabidopsis, potato, tomato, tobacco, cotton, carrot, petunia, sunflower, strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley, sorghum and
10 alfalfa. Cereal plants including rye, barley and wheat may also be transformed with a PERK polypeptide, preferably PERK1.

In a preferred embodiment of the invention, plant tissue cells or cultures which demonstrate PERK activity (or increased PERK activity compared to wild type) are selected and plants are regenerated from these cultures. Methods of regeneration will be apparent to
15 those skilled in the art (see Examples below, also). These plants may be reproduced, for example by cross pollination with a plant that does not have PERK activity. If the plants are self-pollinated, homozygous progeny may be identified from the seeds of these plants, for example, using genetic markers. Seeds obtained from the mature plants resulting from these crossings may be planted, grown to sexual maturity and cross-pollinated or self-pollinated.

20 The nucleic acid molecule is also incorporated in some plant species by breeding methods such as back crossing to create plants homozygous for the PERK nucleic acid molecule.

A plant line homozygous for the PERK nucleic acid molecule may be used as either a male or female parent in a cross with a plant line lacking the PERK nucleic acid molecule to
25 produce a hybrid plant line which is uniformly heterozygous for the nucleic acid molecule. Crosses between plant lines homozygous for the PERK nucleic acid molecule are used to generate hybrid seed homozygous for the resistance nucleic acid molecule.

Fragments/Probes

Preferable fragments include 10 to 50, 50 to 100, 100 to 250, 250 to 500, 500 to 1000,
30 1000 to 1500, or 1500 or more nucleotides of a nucleic acid molecule of the invention. A

fragment may be generated by removing a single nucleotide from a sequence in figure 1 (or a partial sequence thereof). Fragments may or may not encode a polypeptide having PERK activity.

The nucleic acid molecules of the invention (including a fragment of a sequence in figure 1 (or a partial sequence thereof) can be used as probes to detect nucleic acid molecules according to techniques known in the art (for example, see US patent nos. 5,792,851 and 5,851,788). The probes may be used to detect nucleic acid molecules that encode polypeptides similar to the polypeptides of the invention. For example, a probe having at least about 10 bases will hybridize to similar sequences under stringent hybridization conditions (Sambrook et al. 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor). Polypeptide fragments of PERK1 are preferably at least 8 amino acids in length and are useful, for example, as immunogens for raising antibodies that will bind to intact protein (immunogenic fragments). Typically the average length used for synthetic peptides is 8-16, 8 being the minimum, however 12 amino acids is commonly used.

Kits

The invention also includes a kit for conferring increased PERK activity to a plant or a host cell including a nucleic acid molecule of the invention (preferably in a composition of the invention) and preferably reagents for transforming the plant or host cell.

The invention also includes a kit for detecting the presence of PERK nucleic acid molecule (e.g. a molecule in figure 1), comprising at least one probe of the invention. Kits may be prepared according to known techniques, for example, see patent nos. 5,851,788 and 5,750,653.

Antibodies

The invention includes an isolated antibody immunoreactive with a polypeptide of the invention. Antibodies are preferably generated against epitopes of native PERK1 or synthetic peptides of PERK1. The antibody may be labeled with a detectable marker or unlabeled. The antibody is preferably a monoclonal antibody or a polyclonal antibody. PERK antibodies can be employed to screen organisms containing PERK polypeptides. The antibodies are also valuable for immuno-purification of polypeptides from crude extracts.

Examples of the preparation and use of antibodies are provided in US Patent Nos. 5,792,851 and 5,759,788. For other examples of methods of the preparation and uses of

monoclonal antibodies, see US Patent Nos. 5,688,681, 5,688,657, 5,683,693, 5,667,781, 5,665,356, 5,591,628, 5,510,241, 5,503,987, 5,501,988, 5,500,345 and 5,496,705. Examples of the preparation and uses of polyclonal antibodies are disclosed in US Patent Nos. 5,512,282, 4,828,985, 5,225,331 and 5,124,147.

- 5 The invention also includes methods of using the antibodies. For example, the invention includes a method for detecting the presence of a PERK polypeptide such as PERK1, by: a) contacting a sample containing one or more polypeptides with an antibody of the invention under conditions suitable for the binding of the antibody to polypeptides with which it is specifically reactive; b) separating unbound polypeptides from the antibody; and c)
- 10 detecting antibody which remains bound to one or more of the polypeptides in the sample.

Research Tool

- Cell cultures, seeds, plants and plant parts transformed with a nucleic acid molecule of the invention are useful as research tools. For example, one may obtain a plant cell (or a cell line,) that does not express PERK1, insert a PERK1 nucleic acid molecule in the cell, and
- 15 assess the level of PERK1 expression and activity.

The PERK nucleic acid molecules and polypeptides including those in figure 1 are also useful in assays. Assays are useful for identification and development of compounds to inhibit and/or enhance polypeptide function directly.

Suitable assays may be adapted from, for example, US patent no. 5,851,788

20 Using Exogenous Agents in Combination with a Vector

The nucleic acid molecules of the invention may be used with other nucleic acid molecules that relate to plant protection, for example, extensin nucleic acid molecules. Host cells or plants may be transformed with these nucleic acid molecules.

PERK1 ACTIVITY

- 25 We show that PERK1 encodes a protein with kinase activity as its sequence predicts. The bacterially expressed catalytic domain fusion protein of PERK1 is tested for kinase activity (Figure 9c). Furthermore, to ensure that the phosphorylation of the fusion protein was not a result of bacterial kinase activity, a mutated catalytic domain was also generated by site directed mutagenesis which introduced a single base pair substitution of a lysine
- 30 residue to a glutamic acid residue (K→E). This mutation modifies the essential invariant

lysine of subdomain II required for phospho-transfer and renders the kinase inactive. Both the wild-type and the mutated catalytic domains of PERK1 were cloned into the pMAL-c expression system, induced for protein production in the presence of IPTG and purified by affinity chromatography on MBP amylose resin (Figure 9a). Figure 9b is a Western blot to confirm the induction and purification of both the wild-type and mutated catalytic domain fusion proteins using an anti-MBP antibody. The wild-type fusion protein appears to be toxic in bacteria which compromises its inducibility and purification (Figure 9a; lane 2). The mutated fusion protein is induced and purified more efficiently, perhaps due to the fact that it is no longer kinase active (Figure 9a; lane 1).

Figure 9c represents a kinase assay performed on affinity purified wild-type and mutated fusion proteins incubated in the presence of γ -³²PdATP. Detection of a phosphoprotein only in lane 1 provides direct biochemical evidence that the wild-type catalytic domain of PERK1 encodes a functional protein kinase that is capable of autophosphorylation (Lane 1) and that the mutation successfully abolished kinase activity (Lane 2).

Since PERK1 is known to encode a protein with kinase activity, phosphoamino acid analysis was performed to determine the amino acid specificity of its autophosphorylating activity. The results in figure 9d demonstrate that PERK1 is phosphorylated on serine/threonine residues and is therefore a serine/threonine receptor kinase.

Transgenic Plants

An important focus of agricultural biotechnology research is in devising new strategies to combat crop losses due to plant diseases and pests. Advances in the development and improvement of plant transformation techniques has opened up new avenues for generating crops with enhanced resistance against disease and insect attack (i.e. mechanical wounding). Plants can now be engineered to express pathogen derived compounds that disrupt the infection process or alternatively a more desirable and perhaps effective approach is to enhance a plant's endogenous defense mechanisms. Furthermore, the downstream defense mechanisms are fairly well understood, therefore manipulation of some potential upstream signals that control the battery of defense proteins may be a promising strategy to engineer plants with enhanced and broad-spectrum resistance against both insect injury and pathogen attack. PERK1 functions as one such gene.

One approach to show the role of a particular gene in the regulation of gene expression in response to pathogen attack or mechanical wounding has been to generate transgenic plants constitutively expressing the respective cDNA in the sense [1-2] or antisense [3] orientations and examining the effects of wounding and pathogen attack on known downstream target genes involved in these processes. Generating plants (including plant parts and seeds) that overexpress the antisense PERK1 transcript shows the role of PERK1 in mediating a plant's defense response to both wounding and pathogen attack. By mechanically injuring, inoculating with pathogens (i.e. *Sclerotinia sclerotiorum*, *Cylindrosporium conertricum*, *Phoma lingam*) or treating with chemical elicitors and looking at the levels of downstream genes directly involved in these various processes, we directly implicate PERK1 in these pathways. For example, the expression of the functional PERK1 protein is abolished and induction of downstream genes is reduced, so PERK1 is an important upstream component of the pathway. PERK1 offers protection against wounding and pathogen attack, so plants overexpressing PERK1 in the sense orientation exhibit an accumulation of downstream target transcripts involved in these responses and ultimately enhanced resistance.

We score for a phenotype such as enhanced survival of plants overexpressing PERK1 compared to the enhanced susceptibility of plants overexpressing antisense PERK1 transcripts in response to pathogen treatment.

(References 1. Tang, X. et al. (1999). Overexpression of *Pto* activates defense responses and confers broad resistance. *Plant Cell* 11, 15-29; 2. Cao, H. et al. (1998). Generation of broad-spectrum disease resistance by overexpression of an essential regulatory gene in systemic acquired resistance. *Proc. Natl. Acad. Sci.* 95, 6531-6536; 3. Royo, J. et al. (1999). Antisense-mediated depletion of a potato lipoxygenase reduces wound induction of proteinase inhibitors and increases weight gain of insect pests. *Proc. Natl. Acad. Sci.* 96, 1146-1151.)

Transgenic PERK1 *Brassica napus* Plants

In order to generate transgenic *Brassica napus* plants expressing altered levels of PERK1, the full length PERK1 cDNA is cloned into a plant transformation vector in the sense orientation downstream of the constitutively active 35S promoter from the cauliflower mosaic virus (35S CaMV). Subsequently, this expression construct is introduced into an *Agrobacterium tumefaciens* strain in order to transform *B. napus* plants via *Agrobacterium*

mediated transformation. Plant transformants containing the integrated PERK1 cDNA are selected and tested for the presence of the transgene.

Upon generating transgenic plants constitutively expressing the PERK1 cDNA in the sense orientation, we examine the effects of wounding and pathogen attack on known downstream genes involved in these processes to confirm the involvement of PERK1 in these pathways. PERK1 is an upstream component of these pathways. It is involved in mediating a plant's defense response to both wounding and pathogen attack, and transgenic plants overexpressing the PERK1 cDNA in the sense orientation exhibit an increase in the expression of downstream target genes of these pathways. Furthermore, these plants exhibit an enhanced survival relative to a wild type plant in response to wounding and pathogen attack.

PERK1 is a signaling molecule in response to wounding and pathogen attack by generating transgenic plants expressing an HA-epitope tagged PERK1 protein at either the C or N terminus of the protein. Induction of PERK1 with respect to changes in levels of phosphorylation shows the role of PERK1 in these processes. The PERK1 protein is immunoprecipitated using the anti-HA antibody from total protein extracts prepared from *B. napus* tissue after treatment with a specific stimulus. This approach provides quantitative results for the levels of PERK1 phosphorylation in response to various stimuli.

EXAMPLES

Isolation and Sequence Analyses of PERK1 cDNA

In order to isolate novel receptor-like protein kinases in *B. napus* a combination of degenerate oligonucleotide primers designed against conserved kinase subdomains I and VII (Hanks and Quinn, 1991) were used to amplify mass excised phagemid DNA from a newly constructed lambda-pistil cDNA library. The cDNAs encoding products of the expected length (~420-450 bp) were cloned and the deduced partial sequences were analysed against several databases in order to determine which clone represented a novel kinase. One of several candidates, showed the highest degree of sequence similarity to the cytoplasmic domain of known plant receptor protein kinases, and was therefore used to screen the amplified lambda-pistil cDNA library. Several positive clones obtained from the library screen were completely sequenced and a partial 1512 bp consensus sequence was generated to represent the PERK1 cDNA isolated from the library screen. Although this partial PERK1

cDNA had an open reading frame, it did not encode a full length transcript, therefore the 5' end was completed by 5' RACE (see Methods).

The deduced amino acid sequence of PERK1 is shown in Figure 1A and a schematic representation of its hydropathy plot is shown in Figure 1B. The full length cDNA sequence is 2189 bp and consists of one large open reading frame of 1944 bp encoding a predicted protein of 648 amino acids with an estimated molecular mass of 69 kDa (Figure 1). The first methionine of this open reading frame is preceded by two in frame stop codons, TAA and TGA at positions -48 to -45 and -23 to -21 respectively. In addition, there is also an AGAA sequence at position -9 to -6 (Figure 1) which is a favourable site for translation initiation in all eukaryotes (Lutcke et al., 1987).

PERK1 encodes a receptor-like kinase possessing an extracellular domain, a single membrane spanning domain and an intracellular kinase domain (Figure 1B) with four potential N-linked glycosylation sites (Asn-X-Ser/Thr) found throughout the sequence (Figure 1A) (Weinstein et al., 1982). The predicted polypeptide sequence was analyzed using the PSORT database and determined to be a Type Ib intergral membrane protein with a hydrophilic amino terminal domain exposed on the exterior of the membrane but whose coding sequence does not indicate a cleavable signal sequence preceding this domain. Singer (1990) proposes that despite the lack of a signal peptide, Type Ib integral membrane proteins are inserted into the membrane via the usual ER-translocator protein machinery with some slight modifications. The extracellular domain of this protein consists of 137 amino acids (Figure 1A) rich in proline and sharing sequence similarity with extensins, a family of hydroxyproline-rich glycoproteins (HRGPs) that constitute a major protein component of higher plant cell walls (Showalter, 1993). Extensin proteins have two proposed functions in plants, one which contributes to the structural support of the cell wall by forming glycoprotein networks and the other which involves plant defense; helping to protect the plant against mechanical wounding or pathogen attack (Wilson and Fry, 1986; Showlater, 1993). A distinctive characteristic prevalent among dicot extensins is the repetitive Ser-(Pro)₄ pentapeptide consensus motif (Showalter, 1993). A unique feature of PERK1's extracellular domain is the predominance of a slightly modified Ser-(Pro)₂₋₃ motif in addition to the presence of one signature pentapeptide block (Figure 1A). In order to investigate the phylogenetic status of PERK1, sixty four deduced amino sequences corresponding to the extracellular and transmembrane domains of extensin, proline rich and other cell wall

proteins were retrieved from Genbank and used to construct a phylogenetic tree (Clustal X). Results from the phylogenetic analysis indicated that PERK1 is most similar to a subset of extensin proteins as shown by the sequence homology restricted predominantly to the serine/proline rich regions of the protein (data not shown). Extensins and proline rich proteins
5 comprise two major classes of cell wall proteins and are essential for maintaining the proper architecture of a plant cell wall as well as important in helping protect plant cells against wounding and pathogen invasion. These protein families have been the focus of many research efforts and members belonging to these protein classes have been isolated in a wide range of plant systems. A certain degree of sequence identity or conservation is retained
10 among these proteins isolated from different plants. Given that PERK1 shows homology in its extracellular domain to the extensin family of cell wall proteins and is rich in proline residues, homologues of PERK1 exist in many other plant systems.

The protein also contains two other domains of note. Hydropathy analysis (Kyte and Doolittle, 1982) of the protein sequence predicted a membrane spanning region of 23 amino
15 acids (Figure 1; residues 138-160) followed by a characteristic stop transfer sequence rich in charged amino acids [Arg-Arg-Arg] required for the proper insertion in the membrane (Weinstein et al., 1982).

All known protein kinases display amino acid sequence similarity in their catalytic domains which are comprised of eleven subdomains containing some invariant residues
20 important for catalysis (Hanks and Quinn, 1991). The overall features of this organization are identified in the catalytic domain of the PERK1 protein in that all of the absolutely conserved amino acids as well as the highly conserved amino acid groups are present (Figure 1A). The sequences of DIKASN in subdomain VI and GTFGYLAPE in subdomain VIII (Hanks and Quinn, 1991) are strong indicators that PERK1 may possess serine/threonine
25 rather than tyrosine substrate specificity (Figure 1A).

PERK1 is a Single Copy Gene and Ubiquitously Expressed in *B. napus* Tissue

As shown in Figure 2, Southern blot analysis was performed under conditions of varying stringency using *B. napus* genomic DNA digested with several restriction enzymes in order to determine copy number of PERK1 in the Brassica genome. Based on known restriction sites
30 within the cDNA and identical hybridization patterns obtained for low and high stringency conditions (Figure 2), PERK1 exists as a single copy gene and is not a member of a multigene family.

In order to determine whether PERK1 is expressed in plant tissues, RNA gel blot analysis was performed using total RNA isolated from a variety of *B. napus* tissues as shown in Figure 3. The full length PERK1 cDNA probe used in this experiment detected a transcript of 2.2 kb (Figure 3; upper panel) which is consistent with the size of the full length PERK1 cDNA. The 2.2 kb PERK1 transcript was most abundant in *B. napus*, stem petal and pistil tissue (Figure 3; upper panel). Levels of PERK1 mRNA were also detected in root, leaf and anther tissues albeit at much lower levels (Figure 3; upper panel). As an internal control the blot was reprobed with a 18S rRNA to ensure even loading of the total RNA. A transcript detected with relatively the same intensity in all tissues indicates that equal amounts of total RNA was used (Figure 3; lower panel). The difference in the intensity of the 18S rRNA signal in anther tissue (Figure 3; lower panel) is a common problem associated with the desiccate nature of this tissue.

Changes in PERK1 mRNA in Response to Wounding and Chemical Elicitors

In order to examine whether PERK1 expression could be influenced by external stimuli, leaf and stem tissue of *B. napus* plants were wounded and the abundance of PERK1 mRNA was determined by standard Northern blot analysis using the full length PERK1 cDNA as a probe (see Methods). Figure 4 shows changes in the steady-state levels of PERK1 mRNA accumulation following injury. PERK1 transcripts in wounded leaf tissue began to accumulate 5 min after wounding, reaching maximal levels within 15 min post injury represented by an 12fold induction. A 4.5 fold increase in PERK1 mRNA levels was detected 45 min following treatment declining towards basal levels by 2 hr (Figure 4).

A similar profile of PERK1 mRNA steady state levels was obtained for wounded stem tissue (Figure 4). An accumulation of PERK1 mRNA in stem is evident 5 min following wounding which represents a 3.6 fold induction of this gene. Maximum steady state levels of PERK1 mRNA in stem was achieved 30 min after injury corresponding to a 7 fold induction. PERK1 mRNA levels also accumulate in wounded leaf disc tissue as well as following an abrasive wounding treatment (Figure 4a, 4b). Levels of PERK1 mRNA also increase rapidly in the roots of hydroponically grown *Brassica napus* plants (Figure 4c). Therefore, the overall kinetics of PERK1 mRNA accumulation in leaf, stem and root tissue after mechanical wounding is clearly a very rapid response (Figure 4).

Defense mechanisms deployed by plants in response to wounding or pathogen attack have been shown to be induced by certain plant derived chemicals such as methyl jasmonate

(MeJA) and salicylic acid (SA). In order to examine changes in the levels of PERK1 mRNA abundance in response to exogenous application of MeJA, *B. napus* plants were thoroughly sprayed with a 5 micromolar MeJA solution. Leaf and stem tissue was subsequently harvested at various times and the steady state levels of PERK1 mRNA were analysed. Figure 5A shows the RNA gel blot and corrected PERK1 mRNA profile for treated leaf tissue during which no significant accumulation of PERK1 mRNA was detected. In response to MeJA, levels of PERK1 transcript in leaf tissue were very weak resembling basal levels in untreated tissue (Figure 3; upper panel). Exogenous application of MeJA to stem tissue had no effect on the accumulation of PERK1 mRNA as shown by the corrected profile in which the fold induction of PERK1 did not deviate substantially from the untreated control (0 hr) (Figure 5B). Furthermore, no increase in the steady state levels of PERK1 mRNA was detected in the appropriate control treatment (0.1% [v/v] ethanol, solvent control for MeJA) at time 0 hr.

Many genes isolated to date that are induced by a pathogenic stimulus can be at least partially induced by SA (Ward et al., 1991). In order to address the potential role of PERK1 in a plant's defense response against pathogen attack, 4mM SA was used as a chemical elicitor and sprayed onto *B. napus* plants. Figure 6A shows that when SA is exogenously applied to leaf tissue, PERK1 mRNA accumulates 15 min following treatment reaching a maximum 5 fold induction 45 min post-treatment. Steady state levels of PERK1 mRNA in treated stem tissue peaked at 45 min corresponding to an approximate 2 fold induction in response to 4mM SA (Figure 6B).

In order to address whether levels of PERK1 mRNA accumulate in response to a pathogen attack, *Brassica napus* leaf tissue was inoculated with the fungal pathogen *Sclerotinia sclerotiorum* (Figure 7). A 2.5 fold induction in the levels of PERK1 mRNA was detected 10hrs following inoculation with the fungus. This evidence suggests that PERK1 may play a role in mediating a plant's defense response against pathogen attack.

MATERIALS AND METHODS

Construction of Lambda-Pistil cDNA Library

Pistils were collected from floral buds of Westar and W1 cultivars 1-2 days before anthesis.

Total RNA was isolated using the method described by Jones et al. (1985), and enriched for poly(A)⁺ mRNA by affinity chromatography using pre-packed oligo (dT)₂₅-cellulose beads

(New England Biolabs, Beverly, MA). Approximately five micrograms of pistil poly(A)⁺ mRNA was used for the construction of a cDNA library using the ZAP-cDNA[®] synthesis kit (Stratagene, La Jolla, CA). The information encoded by the poly(A)⁺ mRNA was reversed transcribed using M-MuLV RT and converted into stable, unidirectional cDNA which was subsequently inserted into a self-replicating Uni-ZAP XR vector, packaged into phage particles in three separate packaging reactions and amplified as described by the manufacturer's procedures (Stratagene, La Jolla, CA). Infection of *Escherichia coli* host strain XLI-Blue yielded a primary library with an average titer of 1.0×10^6 plaque forming units. The primary library was subsequently amplified to obtain an average total of 6.6×10^{10} plaque forming units.

Generation of Novel Receptor-like Protein Kinase Clones

The isolation of novel *Brassica napus* receptor kinases relied upon the newly constructed cDNA library and involved *in vivo* mass excision of the pBluecsript phagemids from the Uni-ZAP XR vectors as outlined by the manufacturer (Stratagene, La Jolla, CA).

Following efficient mass excision, phagemid DNA was extracted using a large scale alkaline protocol as described by Sambrook et al. (1989) and subjected to the polymerase chain reaction (PCR) using two separate oligonucleotide combinations, RK1/RK2 and RK1/RK3 (obtained from M. Cock, École Normale Supérieure de Lyon, France) specifically designed to prime conserved subdomains of the catalytic domain of receptor protein kinases. RK1 (5'-ggiggTTTCggiAT^Tc_{Ag}TiTT^AT^TcAA^Aggg - 3') served as the forward primer and was constructed based upon a conserved amino acid consensus (GGFGIV^F/YKG) within subdomain I of the catalytic domain. The degeneracy of one reverse primer RK2 (5' - AAiATiC^Tg_{ig}CCATiCC^AgAA^Ag^Tc - 3') reflects a conserved amino acid consensus (DFGMARIF) of subdomain VII which closely resembles the SRKs in *Brassica*. The second reverse oligonucleotide RK3 (5' - A^B_AiA^B_A^TcTTigCiA^Ag_iCC^AgAA^AgTC - 3') was generated based upon conserved amino acids (DFGLAKLL) within subdomain VII prevalent among the RLKs isolated in *Arabidopsis*. Phagemid DNA was amplified in a reaction mixture containing 1 microliter of excised phagemid DNA, 10x PCR buffer (100mM Tris-HCl pH8.3, 500mM KCl, 15mM MgCl₂), 10mM deoxyribonucleotide triphosphate mixture, 10 micromolar of each oligonucleotide primer and 0.5 microliter Tsg polymerase (BioBasics, Canada). The PCR reaction was heated at 95°C for 2 min and amplified for 35 cycles under the following amplification conditions: 1 min at 95°C for denaturation, 1 min 30 sec at 50°C

for primer annealing and 1 min at 72°C for synthesis. A final extension cycle of 10 min at 72°C was also incorporated into the amplification program. All PCR products generated of the expected size (420-450 bp) were gel purified, cloned into the pT7Blue plasmid (Novagen, Madison, WI) and introduced into *Escherichia coli* DH5- alpha. Transformants were tested for the presence of an insert and positive clones were sequenced with universal primers (R-20 and U-19) by an ABI automated sequencer (Model 373 STRETCH DNA; Perkin Elmer Corp., Canada Ltd.) using the dideoxychain-terminating method described by Sanger et al. (1977). Sequence analyses performed using DNAsis® software (Hitachi Software, San Bruno, CA) at the nucleotide and amino acid levels.

10 Screening of Lambda-Pistil cDNA Library

The original 351 bp PCR product was used to screen the lambda-pistil cDNA library. Approximately 2×10^6 plaques from the amplified library were screened and plated at a density of 1×10^5 pfu/plate. Duplicate colony lifts were performed according to Sambrook et al. (1989), and prehybridized for 2 hr at 42°C in 50% formamide, 5x Denhardt's solution (1x Denhardt's solution is 0.02% Ficoll, 0.02% DVP, 0.02% BSA), 5x SSC (1x SSC is 0.15M NaCl, 0.015M sodium citrate), 0.1% SDS, 1mM EDTA and 100:g/ml salmon sperm DNA. Filters were subsequently hybridized overnight in the same solution containing the 351 bp PERK1 cDNA radiolabeled by random priming (Feinberg and Vogelstein, 1983) and washed twice with 2x SSC, 0.1% SDS at room temperature for 15 min, followed by two 25 min washes with 0.5x SSC, 0.1% SDS at 55°C. Plaques containing putative positive clones were cored and subjected to several rounds of screening until single isolates representing the PERK1 clone were obtained. Single clone excision to liberate the double stranded pBluescript phagemid was performed on each isolate according to the procedure recommended by the manufacturer (Stratagene, La Jolla, CA). Phagemid DNA digested with EcoRI/XhoI to release the cloned cDNA was subjected to standard plasmid Southern blot analysis as described by Sambrook et al. (1989) and probed with the radiolabeled 351 bp PERK1 cDNA. The membrane was prehybridized at 42°C in 5x SSPE, 10x Denhardt's solution and 0.5% SDS for 2 hr and hybridized overnight at the same temperature in a buffer containing 50% formamide, 5x SSPE and 0.5% SDS. Washing conditions were performed twice at room temperature for 15 min in 2x SSC, 0.1% SDS followed by several 30 min washes at 55-60°C in 0.1x SSC, 0.1% SDS. An intense hybridization signal would confirm whether phagemids isolated from the library screen contained the cloned cDNA of interest.

Several positive clones were sequenced as previously mentioned using both universal and sequence specific primers to generate a consensus sequence representing the PERK1 cDNA clone (1512bp) isolated from the lambda-pistil cDNA library.

Rapid Amplification of cDNA Ends (5'RACE)

5 The 5' end of the PERK1 cDNA was obtained by the procedure for the rapid amplification of cDNA ends originally described by Frohman et al. (1988) using the 5' RACE System, Version 2.0 kit (Gibco-BRL, Gaithersburg, MD). First strand cDNA was synthesized from approximately 300:g of mixed Westar and W1 pistil total RNA using a gene specific primer GSP1 (5'-TAACCAACAAGAgACA-3') designed to anneal approximately
10 300 bp from the 5' end of the PERK1 cDNA (1512 bp) isolated from the library screen. Following cDNA synthesis, the first strand product was purified from unincorporated dNTPs and GSP1 using a GLASS MAX[®] spin cartridge. A homopolymeric tail was added to the 3' end of the cDNA using TdT (terminal deoxynucleotidyl transferase) and dCTP. Tailed cDNA was amplified using a second gene specific primer GSP2 (5'-CCACTCCCAACTTTCAAC -
15 3') designed to anneal 3' to GSP1 with respect to the cDNA, and an abridged anchor primer (Gibco-BRL, Gaithersburg, MD) which annealed to the homopolymeric tail. PCR amplification was carried out for 35 cycles of denaturation at 94°C for 1 min, primer annealing at 55°C for 1 min and extension at 72°C for 2 min, followed by a final extension cycle for 10 min. A PCR product of the expected size (~1 kb) corresponding to the 5' end of
20 PERK1 was gel purified, cloned into the pT7Blue plasmid (Novagen, Madison, WI) and transformed into *Escherichia coli* DH5- alpha. Confirmation of the 5'RACE product was obtained by plasmid Southern blot analysis as described above and by sequential primer based sequencing.

Cloning of Full Length PERK1 cDNA

25 A PCR based approach was used to generate a full length PERK1 cDNA by combining the 5'RACE product cloned into the EcoRV site of pT7Blue with the cDNA isolated from the library screen cloned into the EcoRI/XhoI sites of the pBluescript SK phagemid. A forward primer (5'-ggAAAgCTTgCATgCCTgCAGgTCgAC -3') containing an internal PstI site was designed to anneal upstream to the EcoRV cloning site of pT7Blue. A reverse primer (5'-
30 CgCCTgCAGgTAATACgACTCACTATAggg -3') also containing a PstI site was designed based on pBluescript phagemid sequence immediately 3' to the EcoRI/XhoI cloning site. Full length PERK1 cDNA was generated from a 100 microliter PCR reaction containing 1

microliter (~20ng) of each template (cDNA in pT7Blue and pBluescript phagemid), 10x Pfu Buffer (200mM Tris-HCl pH8.8, 100mM (NH₄)₂SO₄, 20mM MgSO₄, 1% Triton[®]X-100, 1mg/mlBSA), 10mM dNTPs, 50pmol forward and reverse primers and 1microliter Pfu polymerase (Gibco-BRL, Gaithersburg, MD). The samples were heated to 94°C for 5 min and amplified for 30 cycles with a denaturing cycle of 1 min, a primer annealing cycle at 53°C for 1 min followed by an extension cycle for 3 min at 72°C. The resulting PCR product of the expected size (~2.2kb) was gel purified and cloned into the PstI restriction site of pBluescript KS (+/-) II. The full length PERK1 cDNA sequence was confirmed by a sequential primer based sequencing approach using both universal and sequence specific primers as previously described. All DNA and protein sequence analysis was performed using the DNAsis[®] Software (Hitachi Software, San Bruno, CA).

Genomic DNA Isolation and Southern Blot Analysis

Genomic DNA was extracted from approximately one gram of young *Brassica napus* leaf tissue according to the method described by Goring et al. (1992b). Approximately 5:g of genomic DNA was digested with several restriction enzymes (BamHI, EcoRI, HindIII, PstI, XbaI, XhoI), fractionated through a 0.8% agarose gel and transferred overnight in 10x SSC onto Zetaprobe membrane (Biorad, Hercules, CA). This was performed in duplicate to test hybridization conditions under low and high stringencies conditions. After drying, the membranes were prewashed in 0.1x SSC, 0.5% SDS for 25 min at 60°C. The membranes were prehybridized and hybridized as previously described for plasmid Southern blots with the inclusion of 10% dextran sulfate and 50 microgram/milliliter salmon sperm DNA in the hybridization buffer. Washing conditions for genomic southern blots varied depending on the stringency tested. One membrane was washed under conditions of low stringency for 15 min at room temperature in 2x SSC, 0.1% SDS followed by second 15 min room temperature wash in 1x SSC, 0.1% SDS and three final washes at 50°C in 1x SSC, 0.1% SDS. The second membrane was washed under conditions of high stringency by lowering the salt concentration to 0.1x SSC, 0.1% SDS and increasing the temperature to 65°C. The ³²P-labeled 1512 bp PERK1 cDNA probe was generated by random priming as described by Feinberg and Vogelstein (1983). Membranes were subjected to autoradiography (XAR-5 film, Kodak) overnight at -80°C.

Isolation and Northern Blot Analysis of Multiple Tissue RNA

Total RNA was extracted from a mixture of Westar and W1 root, stem, leaf, petal, anther and pistil tissue as described by Jones et al. (1985). Approximately 40 micrograms of total RNA was fractionated on a 1.2% formaldehyde gel (Sambrook et al., 1989) and
5 transferred to Zetaprobe membrane (Biorad, Hercules, CA) in 10x SSC. Hybridization and high stringency wash conditions were conducted as previously described for genomic Southern blot analysis. The membrane was subsequently probed with a 18SrRNA as an internal control for even loading of total RNA.

Plant Treatments

10 *B. napus* plants were grown in a growth chamber at 22°C with a 16hr light period followed by an 8hr dark period at 16°C. Experiments were conducted on two month old plants, and all experiments used one plant per time point from which leaf and stem tissue was harvested.

Wounding of leaf material was performed by punching out leaf discs every 1cm
15 around the perimeter of the leaf blade ensuring that the midvein remained intact, and stems were wounded by slicing into 1-3cm segments. The wounded tissues were placed in petri dishes containing filter paper moistened with 20mM sodium phosphate buffer supplemented with 50 microgram/milliliter chloramphenicol to prevent bacterial contamination of the wounded tissue (Shirsat et al., 1996). A control (0 hr) time point for this experiment was
20 performed by incubating unwounded tissue in the sodium phosphate buffer. Wounded leaf and stem tissue was harvested at different times after wounding (0 hr, 5 min, 15 min, 45 min, 1 hr, 4 hr, 12 hr, 24 hr, 36 hr, 48 hr). The procedures for the other wounding treatments assayed are described in detail in the figure legends (Figure 4a-c).

Plants were thoroughly sprayed with 50 micromolar methyl jasmonate (MeJA;
25 Sigma, St. Louis, MO) (Titarenko et al., 1997) and 4mM salicylic acid (SA; Sigma, St. Louis, MO) solutions (Schweizer et al., 1998). Leaf and stem tissue was harvested at various time points (0 hr, 5 min, 15 min, 30 min, 45 min, 1 hr, 2 hr, 4 hr, 12 hr, 24 hr, 36 hr, 48 hr, 72 hr, 96 hr) following SA and MeJA treatments. Control sprays were performed with the carrying solutions in the absence of the chemical inducer. Carrying solutions were 5mM phosphate
30 buffer, pH 7 for SA, and 0.1% [v/v] ethanol for MeJA. The method of inoculation used for the fungal treatment is described in the figure caption for Figure 7.

Total RNA was extracted from treated tissue according to the method described by Cock et al. (1997). Depending on the treatment, varying amounts of total RNA (20-40 micrograms) was electrophoresed on a 1.2% formaldehyde and standard Northern blot analysis was performed as described by Sambrook et al. (1989). Hybridization and washing conditions were performed as outlined for the multiple tissue northern blot. Following autoradiography, the amounts of radioactive signal were quantified using Instant Imager Electronic Autoradiography (Packard, Meriden, CT). The membranes were reprobed with the cyclophilin cDNA and the amounts of hybridized radiolabeled cyclophilin were quantified in the same manner. The relative amounts of RNA hybridized to the full length PERK1 cDNA probe were determined after correction for differences in the amounts of cyclophilin RNA.

The present invention has been described in detail and with particular reference to the preferred embodiments; however, it will be understood by one having ordinary skill in the art that changes can be made thereto without departing from the spirit and scope of the invention. All articles, patents and other documents described in this application (including database sequences and/or accession numbers), US patent application no. 60/149,466 filed on August 19, 1999 (entitled "Characterization of a Novel Receptor Kinase from *Brassica* with a Putative Role in Plant Defence"), US patent application no. 60/159,122 and US patent nos. 5,612,191, 5,763,211, 5,750,848 and 5,681,714, are incorporated by reference in their entirety to the same extent as if each individual publication, patent or document was specifically and individually indicated to be incorporated by reference in its entirety. They are also incorporated to the extent that they supplement, explain, provide a background for, or teach methodology, techniques and/or compositions employed herein.

REFERENCES

- Braun, D.M., and Walker, J.C. (1996). Plant transmembrane receptors: new pieces in the signalling puzzle. *TIBS* **21**, 70-73.
- 5
- Cassab, G. I. (1998). Plant cell wall proteins. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **49**, 281-309.
- Chang, C., Schaller, G.E., Patterson, S.E., Kwok, S.F., Meyerowitz, E.M., and Bleeker, A.B. (1992). The TMK1 gene from *Arabidopsis* codes for a protein with structural and biochemical characteristics of a receptor protein kinase. *Plant Cell* **4**, 1263-1271.
- 10
- Clark, S., Williams, R.W., and Meyerowitz, E.M. (1997). The *CLAVATA1* gene encodes a putative receptor-kinase that controls shoot and floral meristem size in *Arabidopsis*. *Cell* **89**, 575-585.
- 15
- Cock, J.M., Swarup, R., and Dumas, C. (1997). Natural antisense transcripts of the S locus receptor kinase gene and related sequences in *Brassica oleracea*. *Mol. Gen. Genet.* **255**, 514-524.
- 20
- Corbin, D.R., Sauer, N., Lamb, C.J. (1987). Differential regulation of a hydroxyproline-rich glycoprotein gene family in wounded and infected plants. *Mol. Cell Biol.* **7**, 4337-4344.
- 25
- Doares, S.H., Narváez-Vásquez, J., Conconi, A., and Ryan, C.A. (1995). Salicylic acid inhibits synthesis of proteinase inhibitors in tomato leaves induced by systemin and jasmonic acid. *Plant Physiol.* **108**, 1741-1746.

Dwyer, K.G., Kandasamy, M.K., Mahosky, D.I. Acciai, J., Kudish, B.I. Miller, J.E., Nasrallah, M.E., and Nasrallah, J.B. (1994). A superfamily of S locus-related sequences in *Arabidopsis*: Diverse structures and expression patterns. *Plant Cell* 6, 1829-1843.

- 5 **Feinberg, A.P. and Vogelstein, B. (1983).** A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132, 6-13.

- Frohman, M.A., Dush, M.K., and Martin, G.R. (1988).** Rapid production of full length DNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* 85, 8998-9002.
- 10

Goring, D.R., and Rothstein, S.J. (1992). The S-locus receptor kinase gene in a self-incompatible *Brassica napus* line encodes a functional serine/threonine kinase. *Plant Cell* 4, 1273-1281.

15

Goring, D.R., Banks, P., Fallis, L., Baszczyński, C.L., Beversdorf, W.D., and Rothstein, S.J. (1992b). Identification of an S-locus glycoprotein allele introgressed form *B. napus* ssp. *oleifera*. *Plant J.* 2, 983-989.

- 20 **Hanks, S.K., and Quinn, A.M. (1991).** Protein kinase catalytic domain sequence database: Identification of conserved features of 1^o structure and classification of family members. *Methods Enzymol.* 200, 38-62.

- Hervé, C., Dabos, P. Galaud, J.-P. Rougé, P., and Lescure, B. (1996).** Characterization of an *Arabidopsis thaliana* gene that defines a new class of putative plant receptor kinases with an extracellular lectin-like domain. *J. Mol. Biol.* 258, 778-788.
- 25

Horn, M.A., and Walker, J.C. (1994). Biochemical properties of the autophosphorylation of RLK5, a receptor-like protein kinase from *Arabidopsis thaliana*. *Biochem. et Biophys. Acta* **1208**, 65-74.

- 5 **Jones, J.D.G., Dunsmuir, P., and Bedbrook, J. (1985).** High level expression of introduced chimeric genes in regenerated transformed plants. *EMBO* **4**, 2411-2418.

Keller, B., and Lamb, C.J. (1989). Specific expression of a novel cell wall hydroxyproline-rich glycoprotein gene in lateral root initiation. *Genes Dev.* **3**, 1639-1646.

10

Kohorn, B.D., Lane, S., and Smith, T.A. (1992). An *Arabidopsis* serine/threonine kinase homologue with an epidermal growth factor repeat selected in yeast for its specificity for a thylakoid membrane protein. *Proc. Natl. Acad. Sci. USA* **89**, 10989-10992.

- 15 **Kyte, J., and Doolittle, R.F. (1982).** A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105-132.

Li, J., and Chroy, J. (1997). A putative leucine rich repeat receptor kinase involved in brassinosteroid signal transduction. *Cell* **90**, 929-938.

20

Lin, H.V., Wang, X.F., Ng-Eaton, E., Weinberg, R.A., and Lodish, H.F. (1992). Expression cloning of TGF-type II receptor, a functional transmembrane serine/threonine kinase. *Cell* **68**, 775-785.

- 25 **Lutcke, H.A., Chow, K.C., Mickel, F.S., Moss, K.A., Kern, H.F., and Scheele, G.A. (1987).** Selection of AUG initiation codons differs in plants and animals. *EMBO J.* **6**, 43-48.

Merkouropoulos, G., Barnett, D.C., and Shirsat, A.H. (1999). The *Arabidopsis* extensin gene is developmentally regulated, is induced by wounding, methyl jasmonate, abscissa acid and salicylic acid, and coded for a protein with unusual motifs. *Planta* **208**, 212-219.

5 **Mu, J.-H., Lee, H.-S., and Kao, T. (1994).** Characterization of a pollen expressed receptor-like kinase gene of *Petunia inflata* and the activity of its encoded kinase. *Plant Cell* **6**, 709-721.

10 **Nasrallah, J.B., and Nasrallah, M.E. (1993).** Pollen-stigma signaling in the sporophytic self-incompatibility response. *Plant Cell* **5**, 1325-1335.

15 **Penninckx, I.A.M., Eggermont, K., Terras, F.R.G., Thomma, B.P. Samblanx, G.W. Buchala, A., Métraux, J.P., Manners, J.M. and Broekaert, W.F. (1996).** Pathogen-induced systemic activation of a plant defensin gene in *Arabidopsis* follows a salicylic acid-independent pathway. *Plant Cell* **8**, 2309-2323.

20 **Peña-Cortés, H., Albercht, T., Prat, S., Weiler, E.W., and Willmitzer, L. (1993).** Aspirin prevents wound-induced gene expression in tomato leaves by jasmonic acid biosynthesis. *Planta* **191**, 123-128.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989). Molecular cloning: A Laboratory Manual, 2nd ed. New York: Cold Spring Harbor Laboratory Press.

25 **Sanger, F., Nicklen, S., Coulson, A.R. (1977).** DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**, 5463-5467.

Sauer, N., Corbin, D.R., Keller, B., and Lamb, C.J. (1990). Cloning and characterization of a wound specific hydroxyproline-rich glycoprotein in *Phaseolus vulgaris*. *Plant Cell Environ.* **13**, 257-266.

Schweizer, P., Buchala, A., Dudker, R., and M9traux, J.P. (1998). Induced systemic resistance in wounded rice plants. *Plant J.* **14**, 475-481.

- 5 Shirsat, A.H., Wieczoreky, D., and Kozibal, P. (1996). A gene for *Brassica napus* extensin is differentially expressed on wounding. *Plant Mol. Biol.* **30**, 1291-1300.

Showalter, A.M., Zhou, J., Rumeau, D., Worst, S.G., and Varner, J.E. (1991). Tomato extensin and extensin-like cDNAs: Structure and expression in response to wounding. *Plant*
10 *Mol. Biol.* **16**, 547-565.

Showlater, A.M. (1993) Structure and function of plant cell wall proteins. *Plant Cell* **5**, 9-23.

- Singer, S. J. (1990). The structure and insertion of integral proteins in the membrane.
15 *Annu. Rev. Cell Biol.* **6**, 247-296.

Song, W.-Y., Wang, G.-L., Chen, L., Kim, H.-S., Pi, L.-Y., Gardner, J., Wang, B.,
Holsten, T., Zhai, W.-X., Zhu, L.-H., Fauquet, C., and Ronald, P.C. (1995). A receptor
kinase-like protein encoded by the rice disease resistance gene *Xa21*. *Science* **270**, 1804-
20 1806.

Stein, J.C., Howlett, B., Boyes, D.C., Nasrallah, M.E., and Nasrallah, J.B. (1991).
Molecular cloning of a putative receptor protein kinase gene encoded at the self-
incompatibility locus of *Brassica oleracea*. *Proc. Natl. Acad. Sci. USA* **88**, 8816-8820.

25

Suzuki, K., and Shinshi, H. (1995). Transient activation and tyrosine phosphorylation of a
protein kinase in tobacco cells treated with a fungal elicitor. *Plant Cell* **7**, 639-647.

Swarup, R., Dumas, D., and Cock, J.M. (1996). A new class of receptor-like protein kinase gene from *Arabidopsis thaliana* possessing a domain with similarity of plant lectin genes. *Plant Physiol.* **111**, 347.

- 5 **ten Dijke, P., Ichijo, H., Franzen, P., Schulz, P., Saras, J., Toyoshima, H., Heldin, C., and Miyazono, K. (1993).** Activin receptor-like kinases: A novel subclass of cell surface receptors with predicted serine/threonine kinase activity. *Oncogene* **8**, 2879-2887.

- 10 **Titarenko, E., Rojo, E., Léon, J., and Sanchez-Serrano, J. (1997).** Jasmonic acid - dependent and -independent signalling pathways control wound-induced gene activation in *Arabidopsis thaliana*. *Plant Physiol.* **115**, 817-826.

- 15 **Tobias, C.M., Howlett, B., and Nasrallah, J.B. (1992).** An *Arabidopsis thaliana* gene with sequence similarity to the S-locus receptor kinase of *Brassica oleracea*. *Plant Physiol.* **99**, 284-290.

- 20 **Torii, K.U., Mitsukawa, N., Oosumi, T., Matsuura, Y., Yokoyama, R., Whittier, R.F. and Komeda, Y. (1996).** The *Arabidopsis ERECTA* gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. *Plant Cell* **8**, 735-746.

Truernit, E., Schmid, J., Epple, P., Illig, J., and Sauer, N. (1996). The sink specific stress-regulated *Arabidopsis STP4* gene: Enhanced expression of a gene encoding a monosaccharide transporter by wounding, elicitors and pathogen challenge. *Plant Cell* **8**, 2169-2182.

- 25 **Ullrich, A., and Schlessinger, J. (1990).** Signal transduction by receptors with tyrosine kinase activity. *Cell* **61**, 203-212.

Usami, S., Banno, H. Ito, Y., Nishihama, R., and Machida, Y. (1995). Cutting activates a 46-kilodalton protein kinase in plants. *Proc. Natl. Acad. Sci. USA* **92**, 8660-8664.

Walker, J.C. (1993). Receptor-like protein kinase genes of *Arabidopsis thaliana*. *Plant J.* **3**, 451-456.

- 5 **Walker, J.C. (1994).** Structure and function of the receptor-like protein kinases of higher plants. *Plant Mol. Biol.* **26**, 1599-1609.

Walker, J.C., and Zhang, R. (1990). Relationship of a putative receptor protein kinase from maize to the S-locus glycoproteins of *Brassica*. *Nature* **345**, 743-746.

10

Wang, G.-L., Ruan, D.-L., Song, W.-Y., Sideris, S., Chen, L., Pi, L.-Y., Zhang, S., Zhang, Z., Fauquet, C., Gaut, B.S., Whalen, M.C., and Ronald, P.C. (1998). Xa21 encodes a receptor-like molecule with a leucine-rich repeat domain that determines race-specific recognition and is subject to adaptive evolution. *Plant Cell* **10**, 765-779.

15

Wang, X., Zafian, P., Choudhary, M., and Lawton, M. (1996). The PR5K receptor protein kinase from *Arabidopsis thaliana* is structurally related to a family of plant defense proteins. *Proc. Natl. Acad. Sci. USA* **93**, 2598-2602.

- 20 **Ward, J.R., Uknes, S.J., Williams, S.C., Dincher, S.S., Wiederhold, D.L., Alexander, D.C., Ahl-Goy, P., Métraux, J.-P., and Ryals, J.A. (1991).** Coordinate gene activity in response to agents that induce systemic acquired resistance. *Plant Cell* **3**, 1085-1094.

Weinstein, J.N., Blumenthal, R., Van Renwoude, J., Kempf, C., and Mauser, R.O. (1982). Charge clusters and the orientation of membrane proteins. *J. Membr. Biol.* **66**, 203-212.

Wilson, L.G., and Fry, J.L. (1986). Extensins: A major cell wall glycoprotein. *Plant Cell Environ.* **9**, 239-260.

Yang, Y., Shah, J., and Klessig, D.F. (1997). Signal perception and transduction in plant
5 defense responses. *Genes Dev.* **11**, 1621-1639.

Zhao, Y., Feng, X.-H., Watson, J.C., Bottino, P.J., and Kung, S.-D. (1994). Molecular
cloning and biochemical characterization of a receptor-like serine/threonine kinase from rice.
Plant Mol. Biol. **26**, 791-803.

10

Zhou, J., Loh, Y.-T., Bressan, R.A., and Martin, G.B. (1995). The tomato gene
Pti1 encodes a serine/threonine kinase that is phosphorylated by Pto and is involved in the
hypersensitive response. *Cell* **83**, 925-935.

We claim:

1. An isolated nucleic acid molecule encoding a proline-rich, extensin-like receptor kinase (PERK) polypeptide, or a fragment of a PERK polypeptide having PERK activity.
2. The molecule of claim 1, wherein the polypeptide is a signaling molecule associated with the cell wall via its extensin-like extracellular domain and is involved in the transduction of extracellular stimuli into an intracellular response through a cytoplasmic kinase domain, thereby bridging the cell wall-plasma membrane continuum.
3. The molecule of claim 1, wherein the extracellular stimuli includes wounding or pathogen attack.
4. The molecule of claim 1, wherein the wounding is selected from the group consisting of a cut, a break, a tear, a fold and an insect wound.
5. The molecule of claim 3, wherein the pathogen comprises bacterial pathogens, fungal pathogens, *Sclerotinia sclerotiorum*, *Cylindrosporium concentricum*, *Phoma lingam*, *Pseudomonas syringae*, *Streptomyces scabies*, Blackleg, Whiterust, *Fusarium* Head Blight, Rust, Bunt, Leaf Spot, White mold, root rot, *Fusarium* ear rot
6. An isolated nucleic acid molecule encoding a PERK polypeptide, a fragment of a PERK polypeptide having PERK activity, or a polypeptide having PERK activity, comprising a nucleic acid molecule selected from the group consisting of:
 - (b) a nucleic acid molecule that hybridizes to a nucleic acid molecule consisting of [SEQ ID NO:1], or a complement thereof under low, moderate or high stringency hybridization conditions wherein the nucleic acid molecule encodes a PERK polypeptide or a polypeptide having PERK activity;
 - (b) a nucleic acid molecule degenerate with respect to (a), wherein the nucleic acid molecule encodes a PERK polypeptide or a polypeptide having PERK activity.
7. The nucleic acid molecule of claim 2, wherein the hybridization conditions comprise low stringency conditions of 1XSSC, 0.1% SDS at 50°C or high stringency conditions of 0.1XSSC, 0.1% SDS at 65°C.

8. An isolated nucleic acid molecule encoding a PERK polypeptide, a fragment of a PERK polypeptide having PERK activity, or a polypeptide having PERK activity, comprising a nucleic acid molecule selected from the group consisting of:

(a) the nucleic acid molecule of the coding strand shown in [SEQ ID NO:1], or a complement thereof;

(b) a nucleic acid molecule encoding the same amino acid sequence as a nucleotide sequence of (a); and

(c) a nucleic acid molecule having at least 17% identity with the nucleotide sequence of (a) and which encodes a PERK polypeptide or a polypeptide having PERK activity

9. The nucleic acid molecule of any of claims 1 to 8, wherein the PERK polypeptide comprises a PERK1 polypeptide.

10. The nucleic acid molecule of claim 1, comprising all or part of a nucleotide sequence shown in [SEQ ID NO:1] or a complement thereof.

11. The nucleic acid molecule of claim 1, consisting of the nucleotide sequence shown in [SEQ ID NO:1] or a complement thereof.

12. A PERK1 nucleic acid molecule isolated from *Brassica*, or a fragment thereof.

13. The nucleic acid molecule of claim 12, wherein the *Brassica* comprises *Brassica napus*, *Brassica juncea*, *Brassica rapa* or *Brassica oleracea*.

14. A recombinant nucleic acid molecule comprising a nucleic acid molecule of any of claims 1 to 4 and a constitutive promoter sequence or an inducible promoter sequence, operatively linked so that the promoter enhances transcription of the nucleic acid molecule in a host cell.

15. The nucleic acid molecule of claim 1, wherein the molecule comprises genomic DNA, cDNA or RNA.

16. The nucleic acid molecule of claim 13, wherein the nucleic acid molecule is chemically synthesized.

17. An isolated nucleic acid molecule comprising a nucleic acid molecule selected from the group consisting of 8 to 10 nucleotides of the nucleic acid molecule of claim 6, 11 to 25

nucleotides of the nucleic acid molecule of claim 6 and 26 to 50 nucleotides of the nucleic acid molecule of claim 10.

18. The nucleic acid molecule of claim 1, comprising at least 30 consecutive nucleotides of [SEQ ID NO:1] or a complement thereof.

5 19. A vector comprising the nucleic acid molecule of any of claims 1 to 4.

20. The vector of claim 19, comprising a promoter selected from the group consisting of a super promoter, a 35S promoter of cauliflower mosaic virus, a chemical inducible promoter, a copper-inducible promoter, a steroid-inducible promoter and a tissue-specific promoter.

10 21. A host cell comprising the recombinant nucleic acid molecule of claim 1 or the vector of claim 19, or progeny of the host cell.

22. The host cell of claim 21, selected from the group consisting of a fungal cell, a yeast cell, a bacterial cell, a microorganism cell and a plant cell.

15 23. A plant, a plant part, a seed, a plant cell or progeny thereof comprising the recombinant nucleic acid molecule of claim 15 or the vector of claim 19.

24. The plant part of claim 23, comprising all or part of a leaf, a flower, a stem, a root or a tuber.

20 25. The plant, plant part, seed or plant cell of claim 23, wherein the plant, plant part, seed or plant cell is of a species selected from the group consisting of *Brassica napus*, *Brassica rapa*, *Brassica juncea*, *Brassica oleracea*, or from the family Brassicaceae, Arabidopsis, potato, tomato, tobacco, cotton, carrot, petunia, sunflower, strawberries, spinach, lettuce, rice, soybean, corn, wheat, rye, barley, sorghum and alfalfa.

26. The plant, plant part, seed or plant cell of claim 23, wherein the plant comprises a dicot plant.

25 27. The plant, plant part, seed or plant cell of claim 23, wherein the plant comprises a monocot plant.

28. An isolated polypeptide encoded by and/or produced from the nucleic acid molecule of any of claims 1 to 4, or the vector of claim 19.

29. An isolated PERK polypeptide or a fragment thereof having PERK activity.

30. An isolated polypeptide, the amino acid sequence of which comprises at least ten consecutive residues of [SEQ ID NO:2].
31. An isolated immunogenic polypeptide, the amino acid sequence of which comprises at least 8 consecutive residues of [SEQ ID NO:2].
- 5 32. An isolated polypeptide, the amino acid sequence of which comprises residues 1 to 137, 138 to 160 and 161 to 648 of [SEQ ID NO:2].
33. The polypeptide of claim 29 comprising all or part of an amino acid sequence in [SEQ ID NO:2].
34. A polypeptide fragment of the PERK polypeptide of claim 29, or a peptide mimetic of the
10 PERK polypeptide.
35. The polypeptide fragment of claim 33, consisting of at least 20 amino acids, which fragment has PERK activity.
36. The fragment or peptide mimetic of claim 34, which is capable of being bound by an antibody to the polypeptide of claim 28.
- 15 37. The polypeptide of claim 28 which is recombinantly produced.
38. An isolated and purified polypeptide comprising the amino acid sequence of a PERK polypeptide, wherein the polypeptide is encoded by a nucleic acid molecule that hybridizes under moderate or stringent conditions to a nucleic acid molecule in [SEQ ID NO:1], a degenerate form thereof or a complement.
- 20 39. A polypeptide comprising a sequence having greater than 20% sequence identity to the polypeptide of claim 33.
40. The polypeptide of claim 28, wherein the polypeptide comprises a PERK polypeptide.
41. The polypeptide of claim 40, isolated from *Brassica*.
42. The polypeptide of claim 41, wherein the *Brassica* comprises *Brassica napus* or *Brassica juncea* or *Brassica rapa* or *Brassica oleracea*.
25
43. The polypeptide of claim 39, comprising a kinase domain including at least 30% homology to the kinase domain of [SEQ ID NO:2] and/or an extracellular domain including at least 20% homology to the extracellular domain of [SEQ ID NO:2].
44. An isolated nucleic acid molecule encoding the polypeptide of claims 29 or 33.

45. An antibody directed against the polypeptide of claim 33.
46. The antibody of claim 45, comprising a monoclonal antibody or a polyclonal antibody.
47. An isolated nucleic acid molecule encoding a polypeptide that reduces the severity of wounding or pathogen attack in a plant, the polypeptide comprising:

- 5 (a) an extracellular domain which recognizes an extracellular binding molecule whose level is increased during the wounding or pathogen attack, the extracellular domain encoding a plurality of repeats selected from the group consisting of SPPPP, SPP, PP and PPP, wherein a plurality of the proline molecules are capable being glycosylated and/or hydroxylated;
- 10 (b) a membrane domain operably connected to the extracellular domain, wherein the membrane domain is capable of extending across a cell membrane from the extracellular side of the membrane to intracellular side of the membrane; and
- (c) a cytoplasmic domain operably connected to the membrane domain, wherein the cytoplasmic domain comprises a means for producing kinase activity when the
- 15 extracellular binding molecule interacts with the extracellular domain.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
1 March 2001 (01.03.2001)

PCT

(10) International Publication Number
WO 01/14563 A1(51) International Patent Classification: C12N 15/52,
C07K 14/415, C12N 15/82, C07K 16/16, A01H 5/00

(21) International Application Number: PCT/CA00/00966

(22) International Filing Date: 18 August 2000 (18.08.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/149,466 19 August 1999 (19.08.1999) US
60/159,122 13 October 1999 (13.10.1999) US(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

(71) Applicants and

(72) Inventors: GORING, Daphne [CA/CA]; 2240 Taylors Orchard, Mississauga, Ontario L5B 2T3 (CA). SILVA, Nancy [CA/CA]; 66 Woodhaven Crescent, Richmond Hill, Ontario L4E 3T2 (CA).

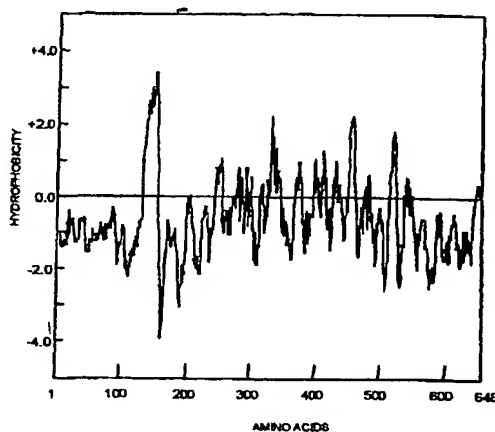
Published:

- With international search report.
- Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

(74) Agent: DEETH WILLIAMS WALL LLP; National Bank Building, Suite 400, 150 York Street, Toronto, Ontario M5H 3S5 (CA).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES



(57) Abstract: The invention includes PERK (Proline-rich Extensin-like Receptor Kinase) nucleic acid molecules and polypeptides. A receptor-like protein kinase designated PERK1 Proline-rich Extensin-like Receptor Kinase 1) was isolated from an 8-pistil cDNA library of *Brassica napus*. The deduced PERK1 protein is comprised of a cytoplasmic domain that contains all of the conserved amino acids prevalent among serine/threonine kinases, a transmembrane domain and an extracellular domain with sequence similarity to the extensin family of plant cell wall proteins. Northern blot analysis demonstrated that PERK1 mRNA accumulated rapidly in stem tissue of *B. napus* in response to wounding and treatment with salicylic acid. In contrast, no significant accumulation of PERK1 mRNA was detected following treatment with methyl jasmonate. The rapidity of PERK1 mRNA accumulation in response to these treatments shows a role in plant defense signaling.

WO 01/14563 A1

A.

ATGTCTCGGCGCCGTCTCCG000 24
 M S S A P S P G 84
 25 ACTGGTTCCGCTCCATCTCCACCATCAAACCTCCACAACCACCACTCTCTCCCAOCTTCC 84
 T O S P P S P P S N' S T T T T P P P A S
 85 GCTCTCTCTCCACCAACCTTCTTCTCTCCGCGCCATCCACTATTCGACATCTCT 144
 A P P P T T P S S P P P P S T I P T S P
 145 CCTCTCTCTCTCGCTCTACACCTTCTGCTCTCTCTCCATCTCCACCAACTCCATCTACO 204
 P P S S R S T P S A P P P S P P T P S T
 205 CCGGATCTCCACCTCTCTCTCTCAGCCGTCTCCACCCGCTCCAACCTACGCCCGOATCT 264
 P O S P P P L P Q P S P P A P T T P Q S
 265 CCACCCGACCTGTGTACTCTCTCTCTCAGAAACCTCCACCTTCAGTCCCAAGGACCCG 324
 P P A P V T P P T R N P P P S V P G P P
 325 TCCAATCTCTCAGCGAAGGATCTCTCTCAGCTCCATCTTCTCCCTCGCGCGCTCT 384
 S N P S R E G O S P R P P S S P S P P S
 385 CCTTCTCTCGACGCTTATCAACAGGAGTGGTGGGAAATCGGCATCGGAGGAGTCTGCT 444
 P S S D G L S T G V V V G I A I G G V A
 445 CTGCTTGTGATAGTGTCTGATTTGTCTCTCTGTAAGAAAGAAACGAGGAGAGAGAA 504
 L L V I V T L I C L L C K K K R R R D E
 505 GAAGATGCTTACTATGTTCTCTCGCCACCTCTCTCTGCTCCAAAGCCGAGGACCTTAC 564
 E D A Y Y V P P P P P P G P K A G O P Y
 565 GGTGGACAGCAGCAACAATGCGGCAACAAACGCAACACCCGTCAGATCATCTCGTG 624
 G G Q Q Q Q W R Q Q N A T P P S D H V V
 625 ACOTCACTACCAACCACTAAGGCTCCATCTCCACCAAGGCAACCTCTCTCCACCTCCA 684
 T S L P P P P K A P S P P R Q P P P P P
 685 CCACCGCTTTCATGAGCAGCAGCGCGCTCCGACTACTCGGACCGTCCAGTCTTCTCT 744
 P P P F M S S S G G S D Y S D R P V L P
 745 CCACCGTCTCCAGGCTTGTGTTAGGCTTCTCCAAAGCACTTTCACATACGAGAGCTA 804
 P P S P G L V L G F S K S T F T Y E E L
 805 GCTAGAGCCACCAATGTTTCTCCGAGCGCAACTTGTAGGACAAGCGGCTTCTGTTAC 864
 A R A T N 'Q F S E D N L L G Q Q Q F Q Y
 865 GTGCACAAAGGTGTGTTGCTAGTGGAAAGAGTGTGTGAAAGCAGTTGAAAGTGGG 924
 V H K Q V L P S 'G K E V A V K Q L K V G
 925 AGTGGTCAGGAGAGAGGAGTTCAGGAGAGGTTGAGATCATCAGCAGATTCACCA 984
 S G 'Q Q E B E F O A E V E I I S R V 'H H
 985 AGGCATCTGTGTCTTGTGTTATGTCATCGCGGTGCCAAAGATTGCTTGTCTAT 1044
 R H L V S L V Q Y C I A 'Q A K R L L V Y
 1045 GAGTTTGTCTTAACAACATCTCGAGCTTCACTCTCCATGCGAGGAGCGGCTTACAATG 1104
 E F V P N N N L E L H L H Q E G R 'P T M
 1105 GAATGGAGCACCAGATTGAAGATTGCTCTTGGATCTGCTAAAGGACTTCTTATCTTCAT 1164
 E W S T R L K I A L G S A K Q L S Y L H
 1165 GAAGATTGCAATCTAAATCATTACCGTGATATCAAGGCTTCAACATATTGATAGAT 1224
 E D C N P K J I L H R D I K A S N I L J P
 1225 TTCAAGTTGAAGCTAAGGTTGCTGATTTTGGCTTCTGCTAAGATTGCTTCTGATACAAAC 1284
 F K E 'E A E V A D F Q L A K I A S D T N
 1285 ACGCATGTATCAACACGTGTGATGGGAACCTTGGGACTTGGCTCCGGAATACGCTGCA 1344
 T H V 'S T R V M Q T F Q Y L A P E X 'A A
 1345 AGCGGAAAGCTCACGAGAGTGTGACGTTTCTCATTTGGCTTGTGCTTTTGGAGCTC 1404
 S Q K L T E K S D V F S F Q V V L L F L
 1405 ATTACTGGACGTGAGCCGTTGATGCAACAATGTCTATGTAGATGACAGCTTAGTGTAC 1464
 I T Q R R P V D A N N 'V Y V Q D S L V D
 1465 TGGGACGACCAATTGCTTAACCGAGCATCTGAGCAAGGAGACTTTGAGGGTTTAGCTGAT 1524
 W A R P L L N R A S E Q Q D F E G L A D
 1525 GCAAGATGAATAATGGGTATGACAGAGAGGAGATGGCTCGCATGGTTGCTTGTGCTCG 1584
 A K M N N G Y D R E 'E M A R N V A C A A
 1585 GCTTGTGTTCCCAATTCAGCTCGCCGAGACCTCGCATGAGCCAGATTGTGCTGCTTA 1644
 A C V R H S A R R R P R M S Q I V R A L
 1645 GAAAGAAATGTATCACTGTCTAGATCTTAAAGGAGGATGAGACCAAGGTCAAAGCAATGTA 1704
 E G N' V S L S D L N E G M R P Q Q S N V
 1705 TACAGCTCATACGAGGAAAGCAGGATTATGACTCGAGCCAATACAATGAAGCATGAAG 1764
 Y S S Y G G S T D Y D S S Q Y N E D M K
 1765 AAGTTTAGGAAATGGCACTTGAACCTCAAGAGTACAACGCCACGGGTGAGTACAGTAAT 1824
 K F R K M A L G T Q E Y N' A T G E Y S N
 1825 CCGACCACTGACTATGACTGTACCCGTCTGGTTCAAGCAGCGAGGCGCAACCAACACGC 1884
 P T S D Y G L Y P S O S S S E G Q T T R
 1885 GAAATGAGATGGGGAAGATTAAAGAGAACCGGTCAAGGTTATAGTGGACCTTCTCTTAA 1944
 E M E M G K I K R T O Q Q Y S O P S L *

Figure 1A

1/26

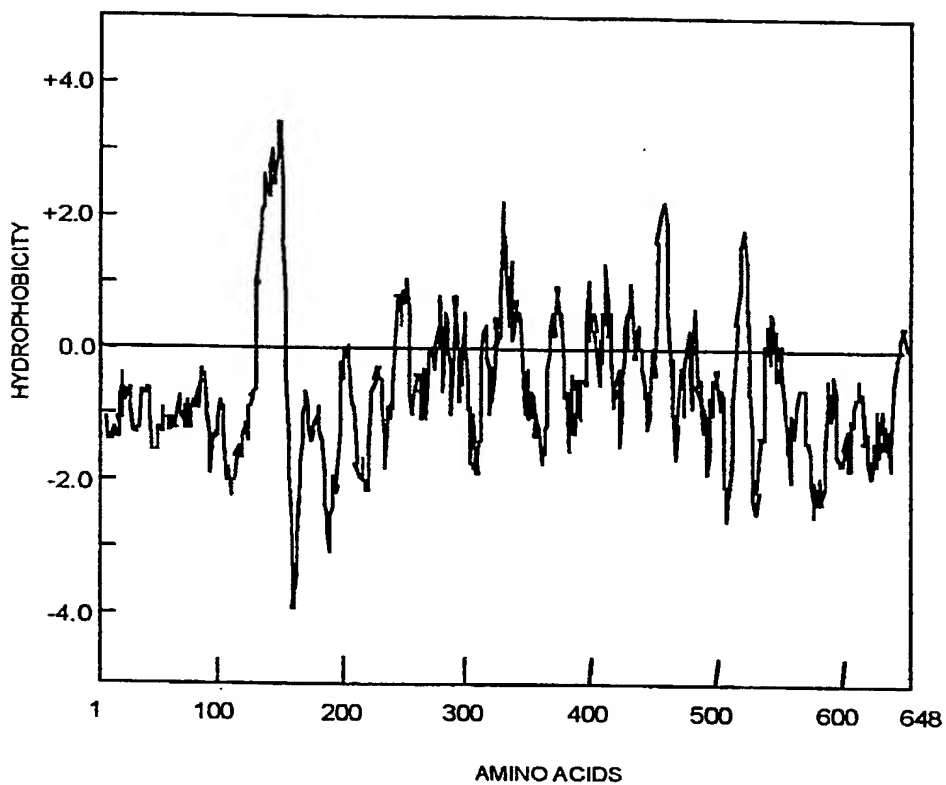
B.

Figure1B
2/26

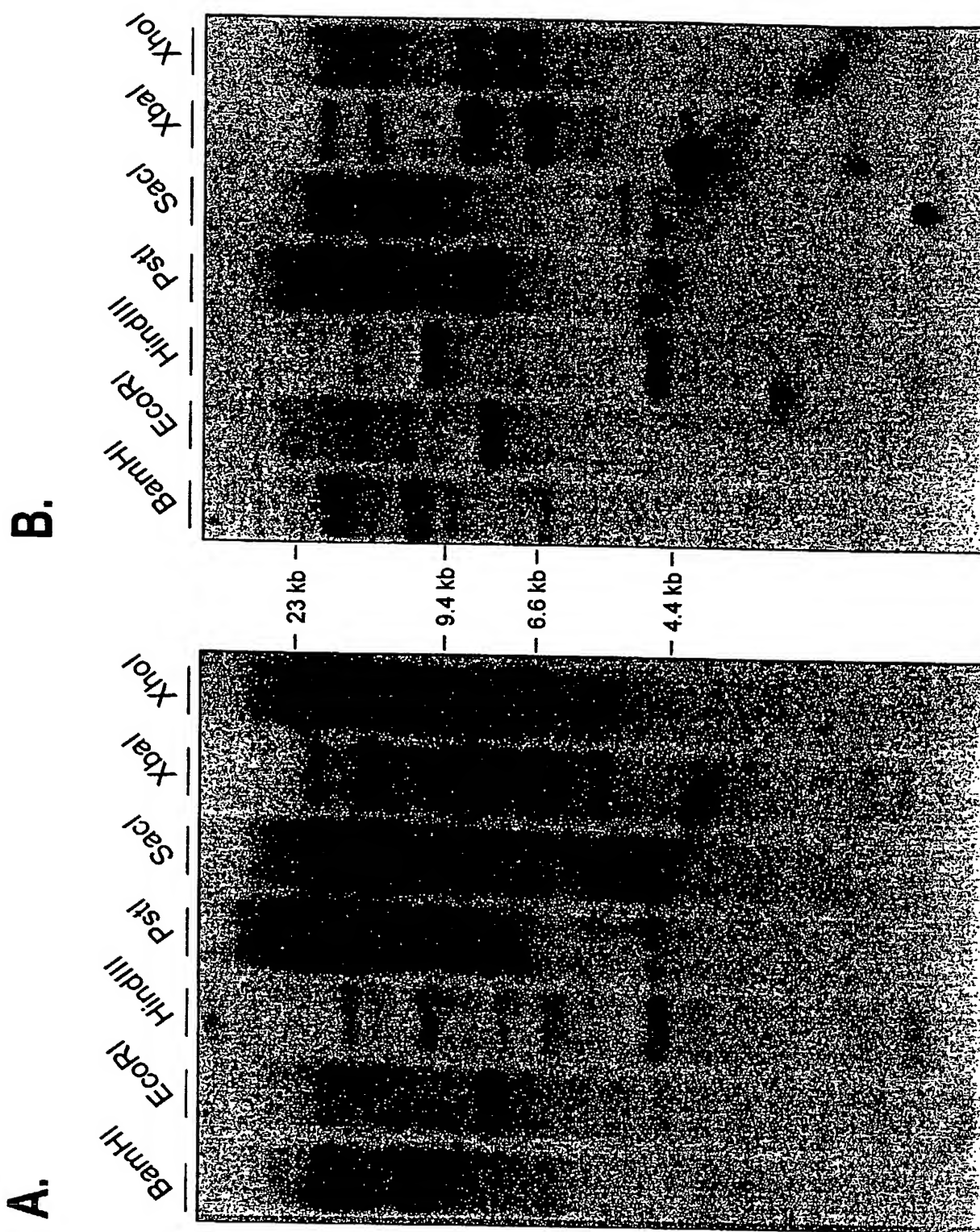


Figure 2

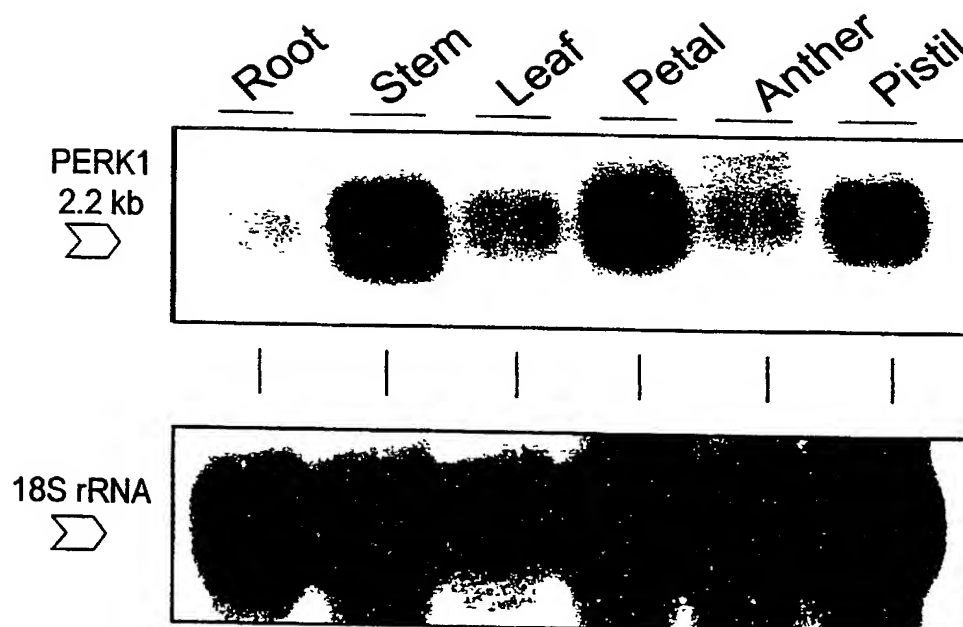
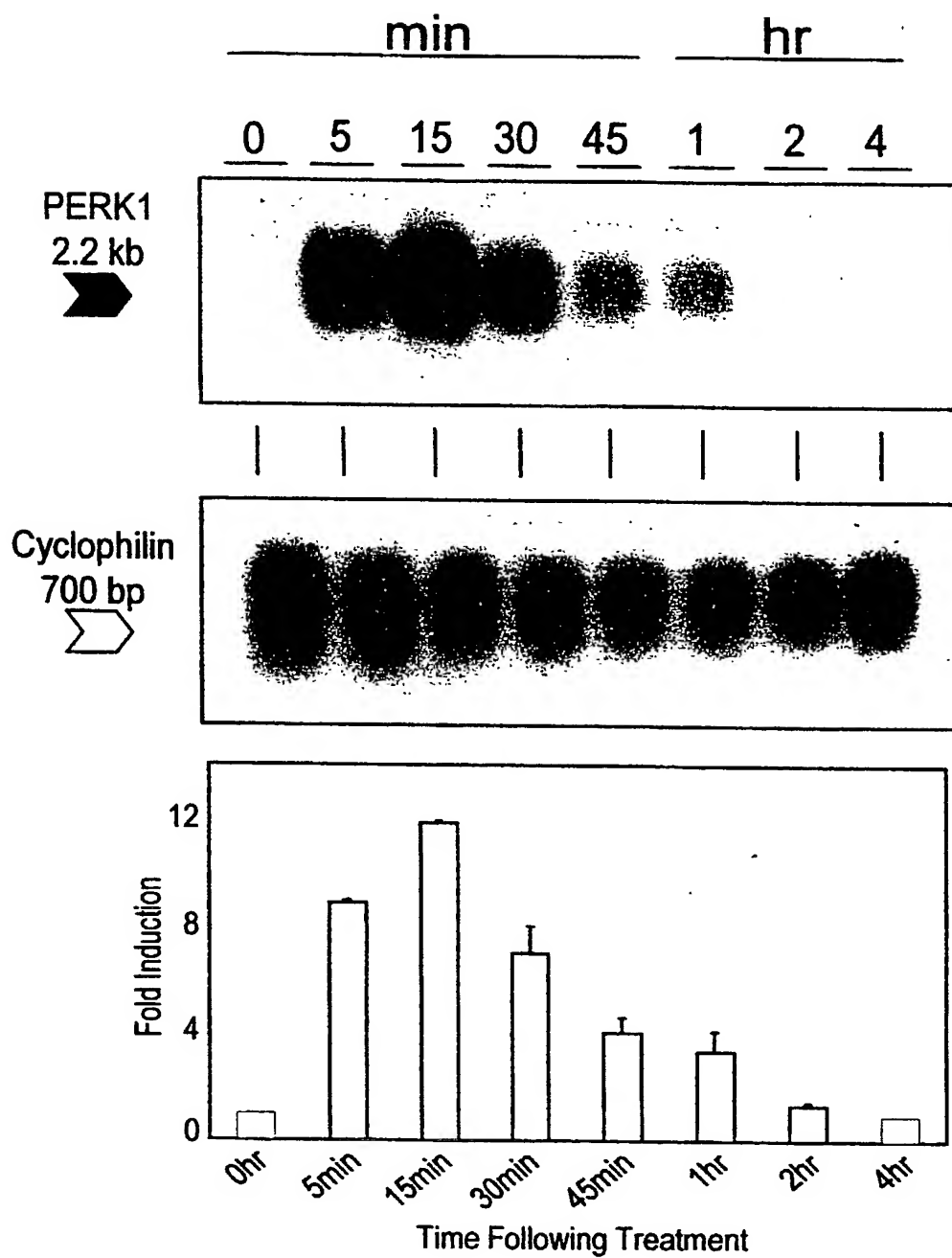


Figure 3
4/26

A.**Figure 4A**

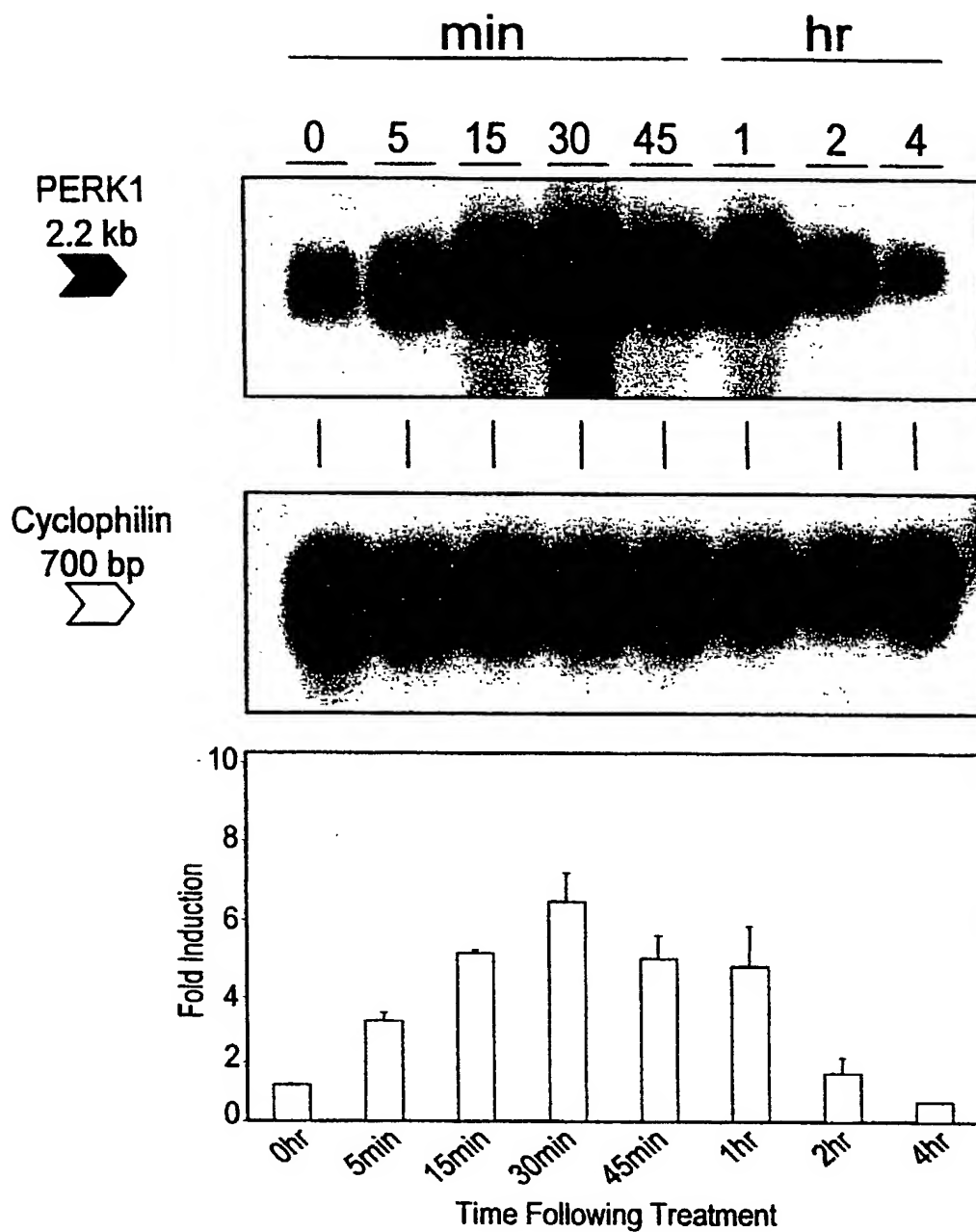
B.

Figure 4B
6/26

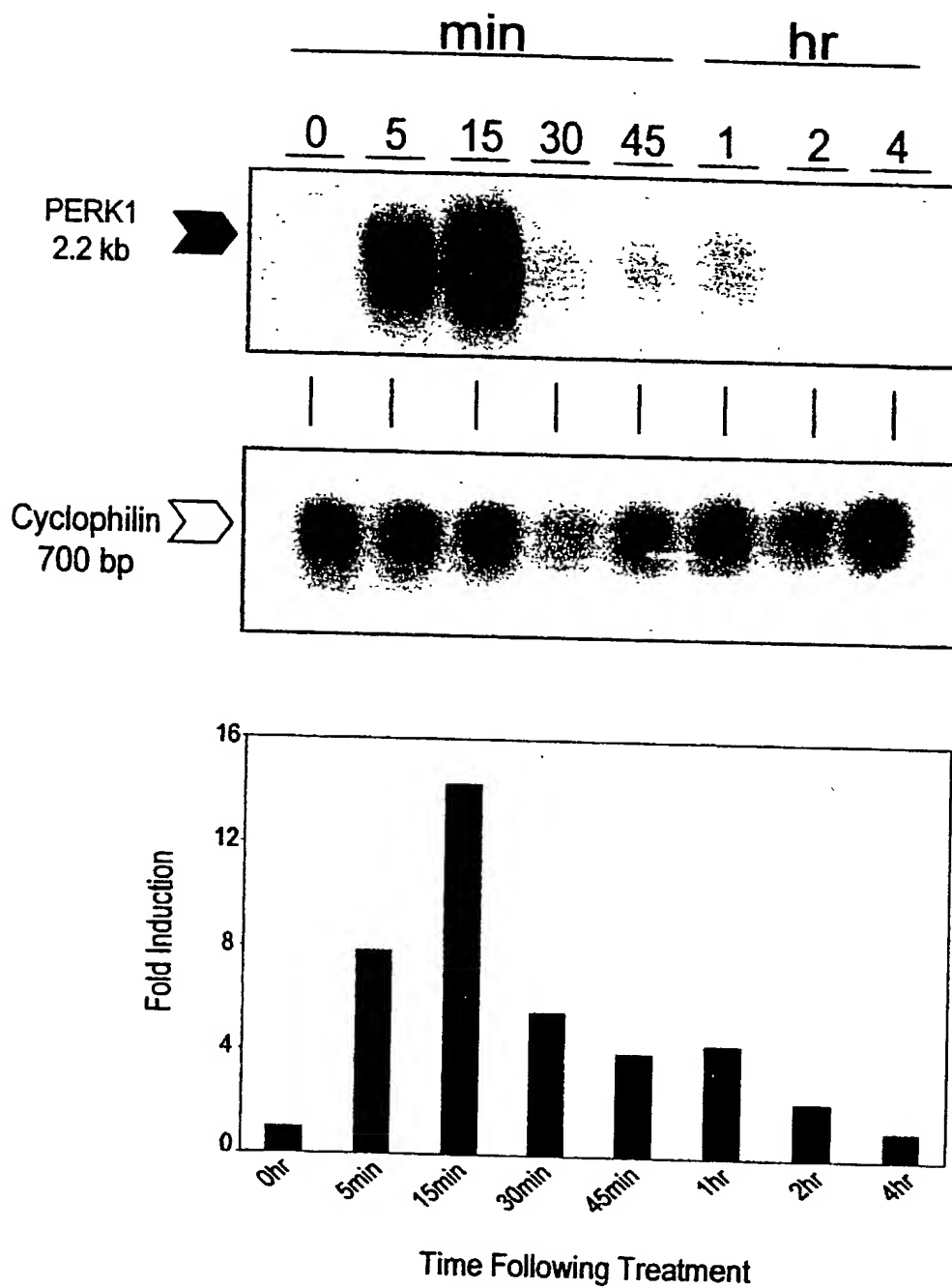


Figure 4a
7/26

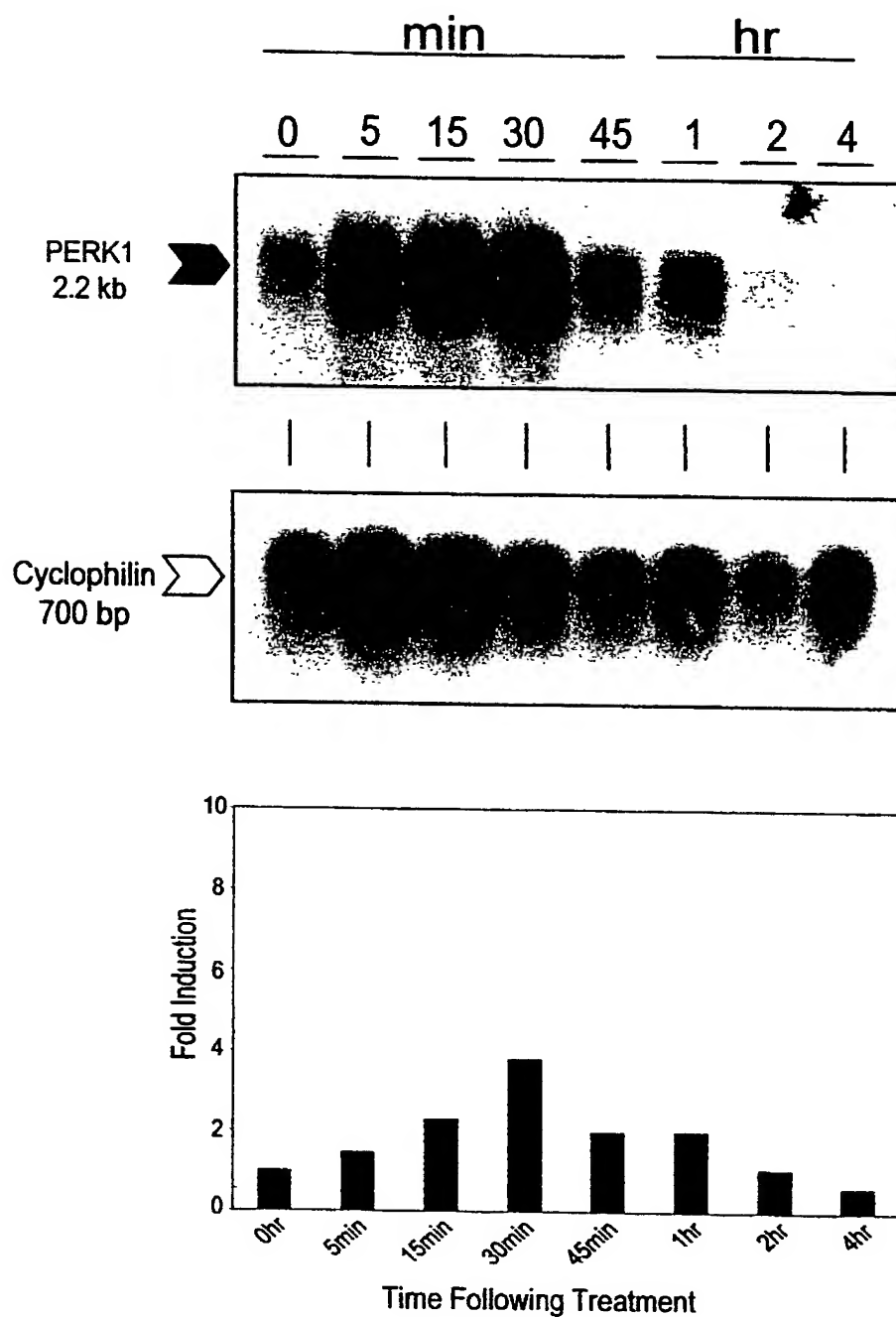
A.

Figure 4b(A)

8/26

SUBSTITUTE SHEET (RULE 26)

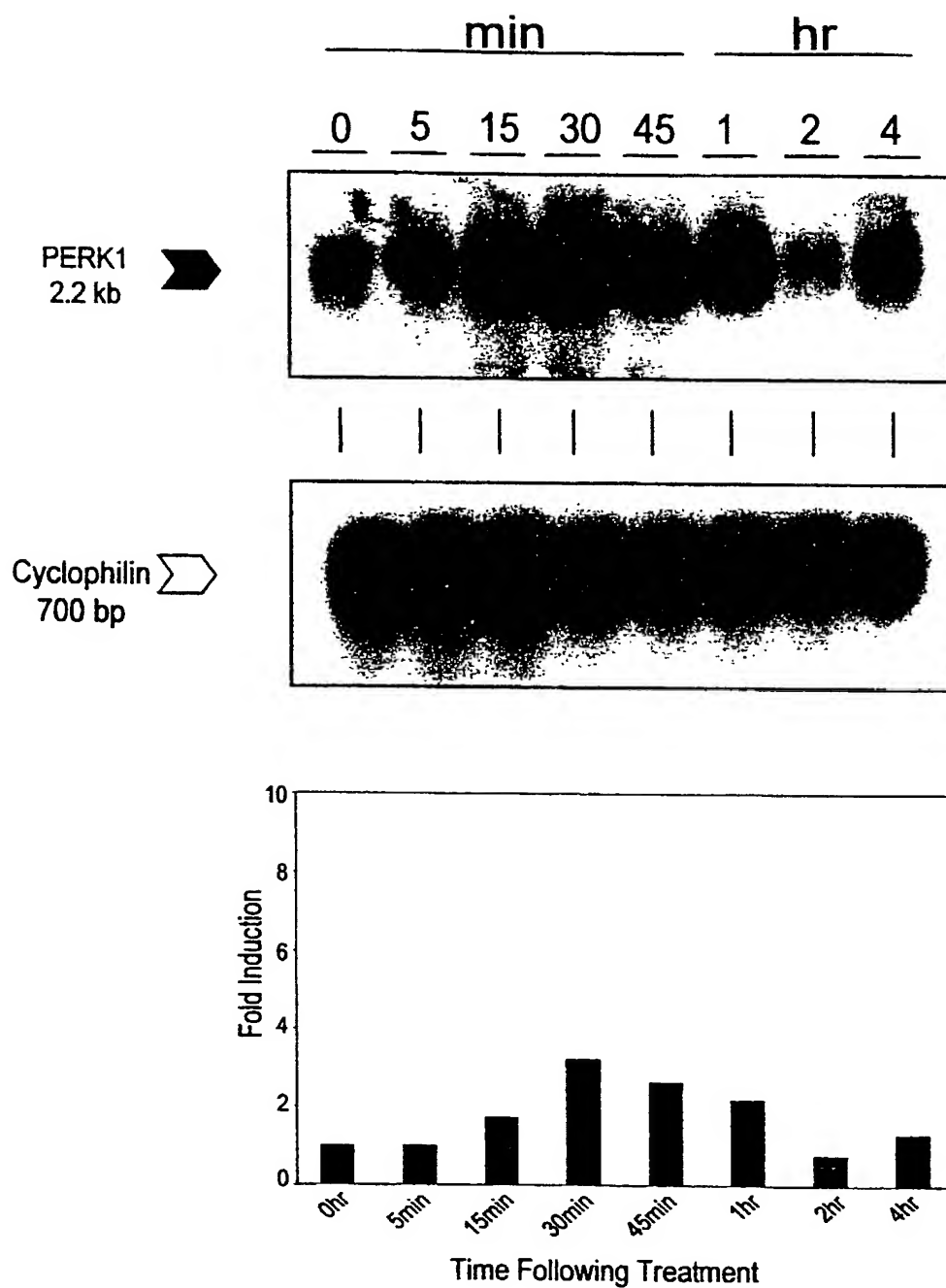
B.

Figure 4b(B)

9/26

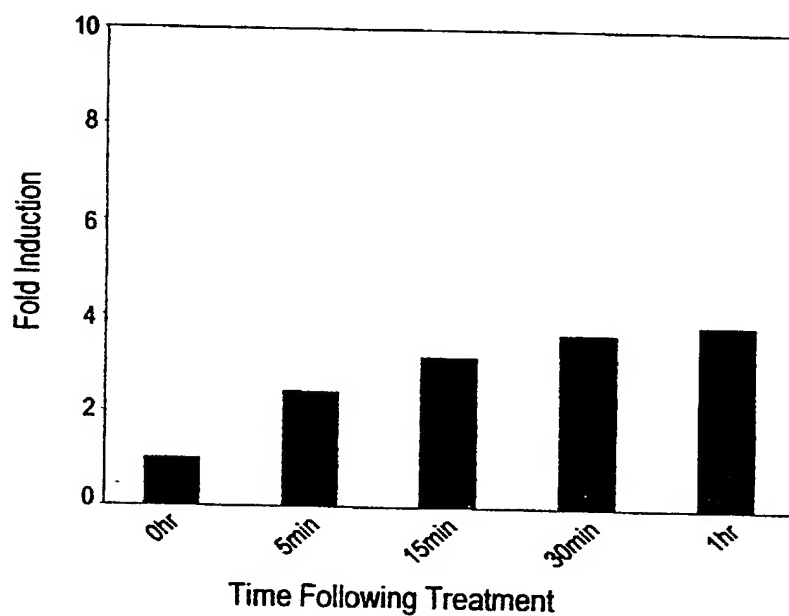
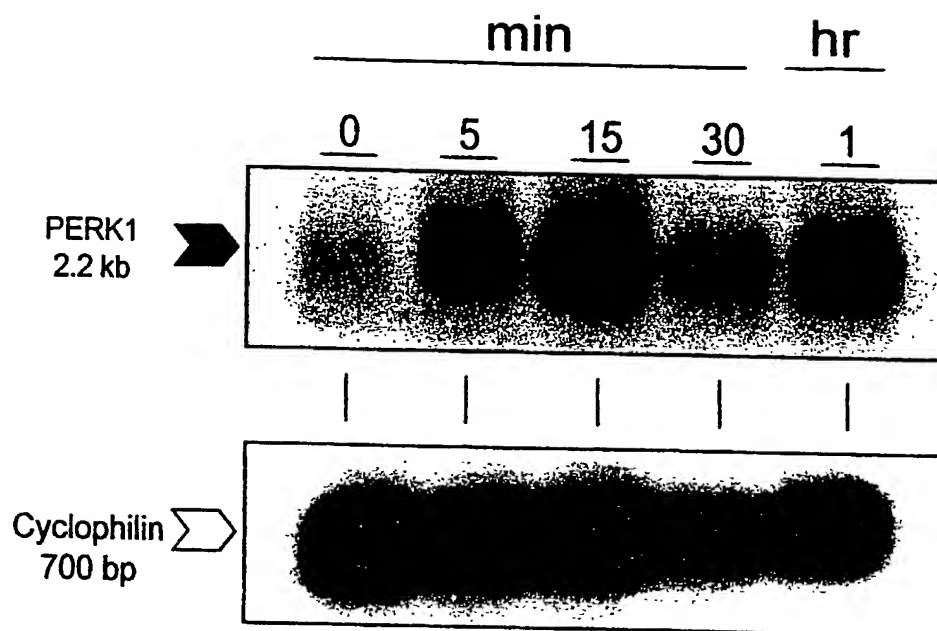


Figure 4c

10/26

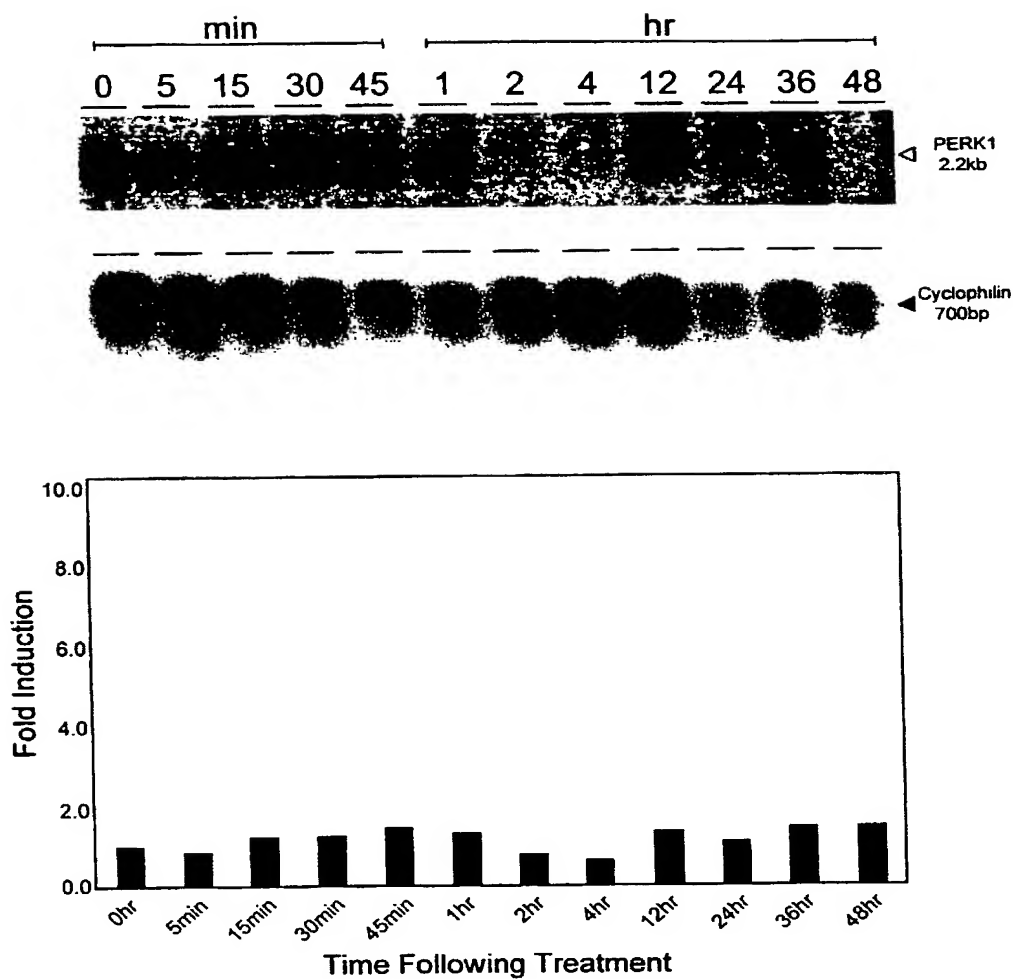
A.

Figure 5A
11/26

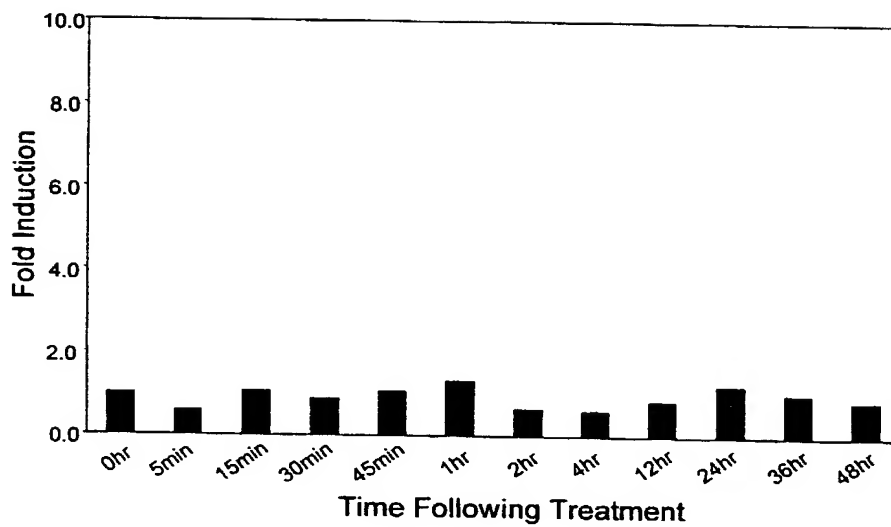
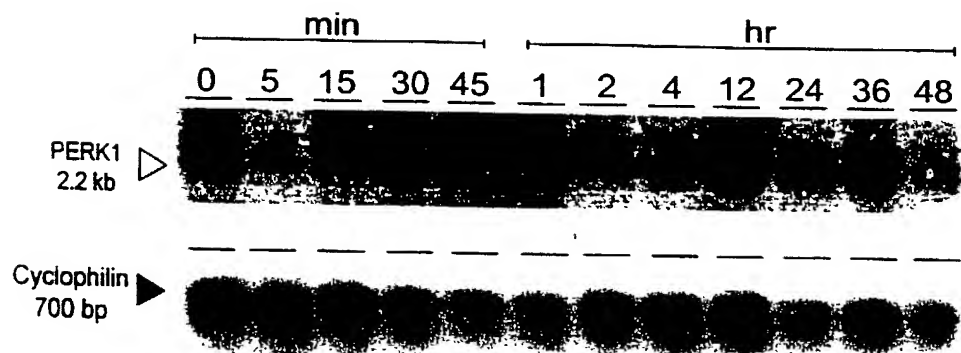
B.

Figure 5B
12/26

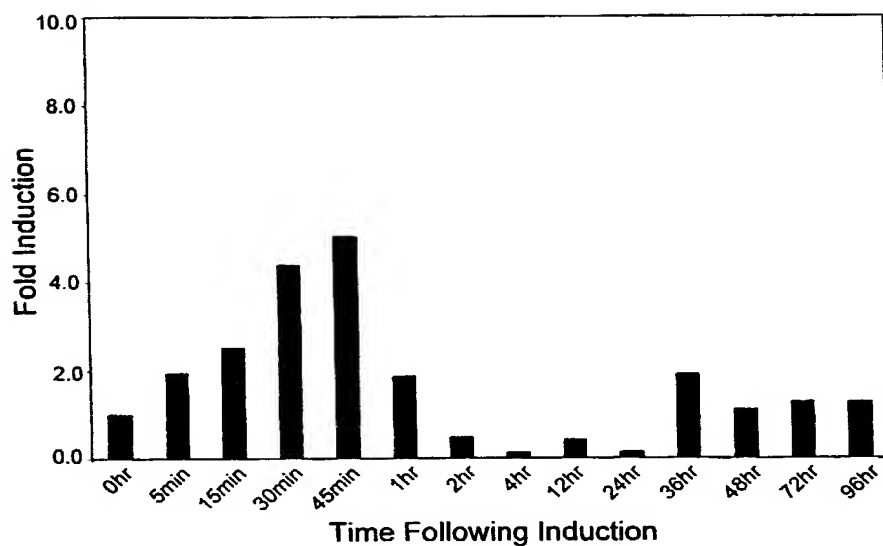
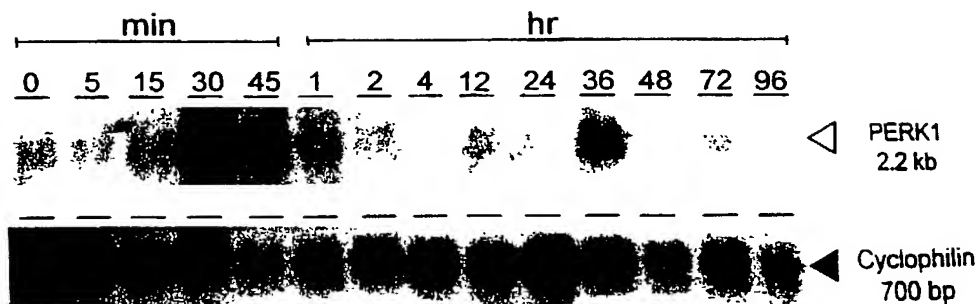
A.

Figure 6A
13/26

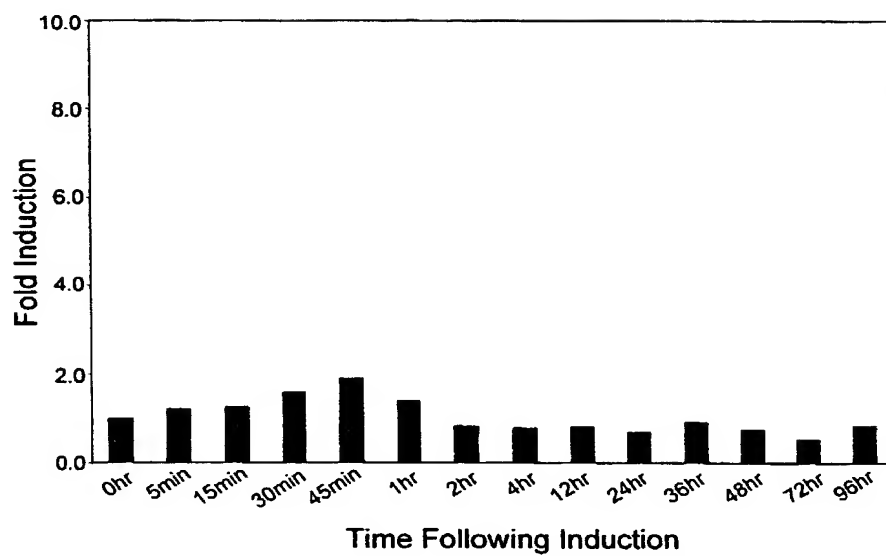
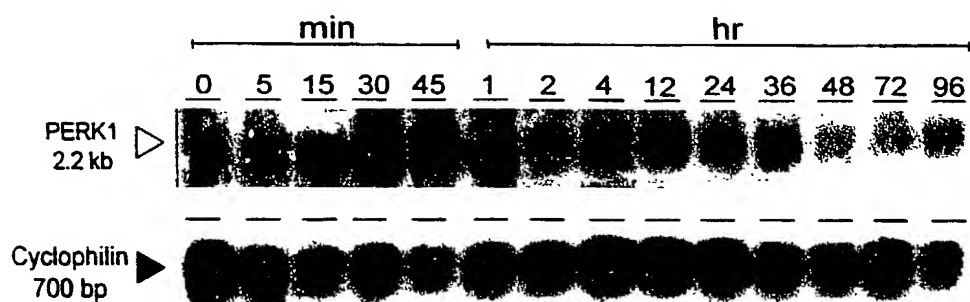
B.

Figure 6B
14/26

- Fungal pathogen : *Sclerotinia sclerotiorum*

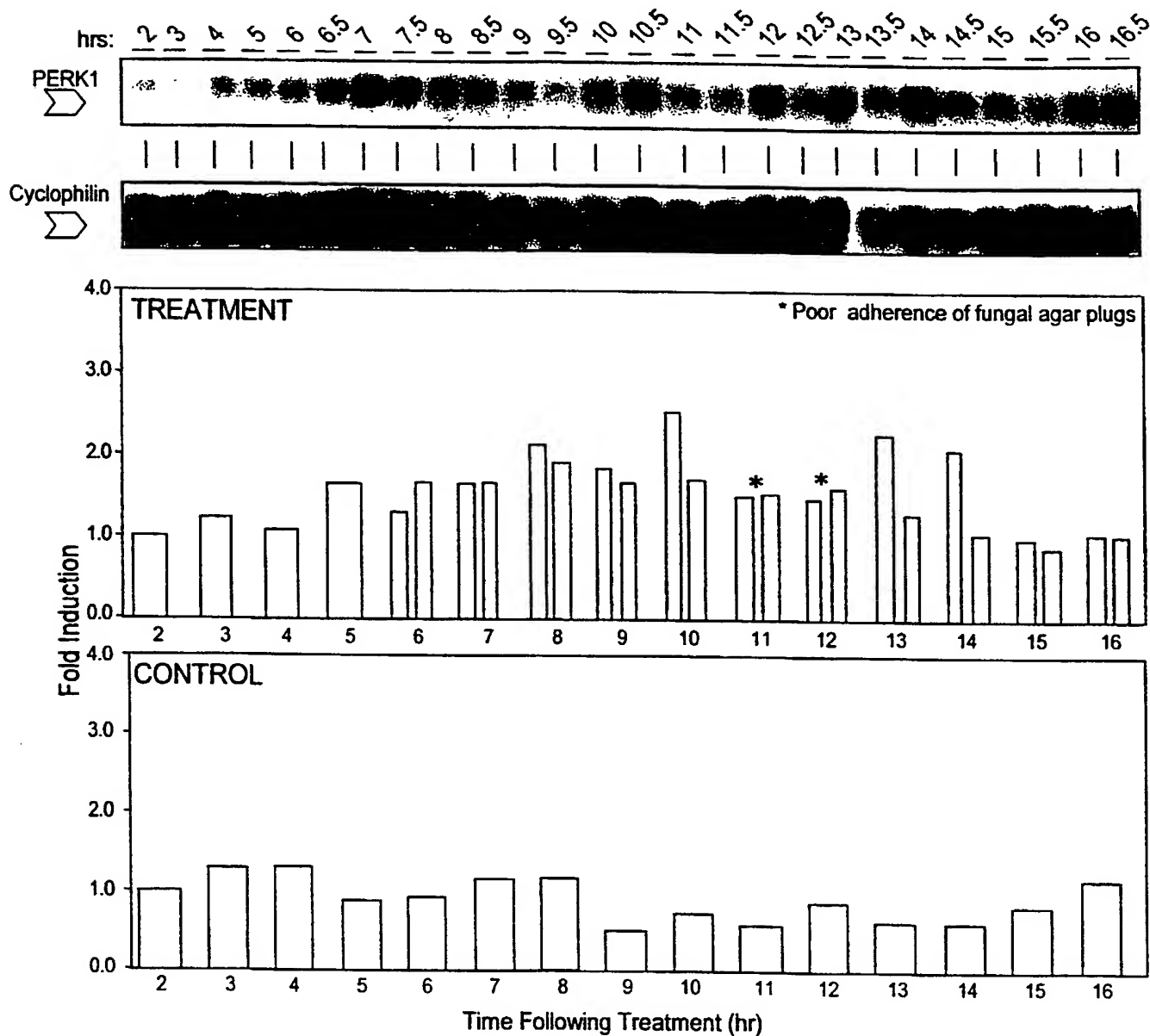


Figure 7
15/26

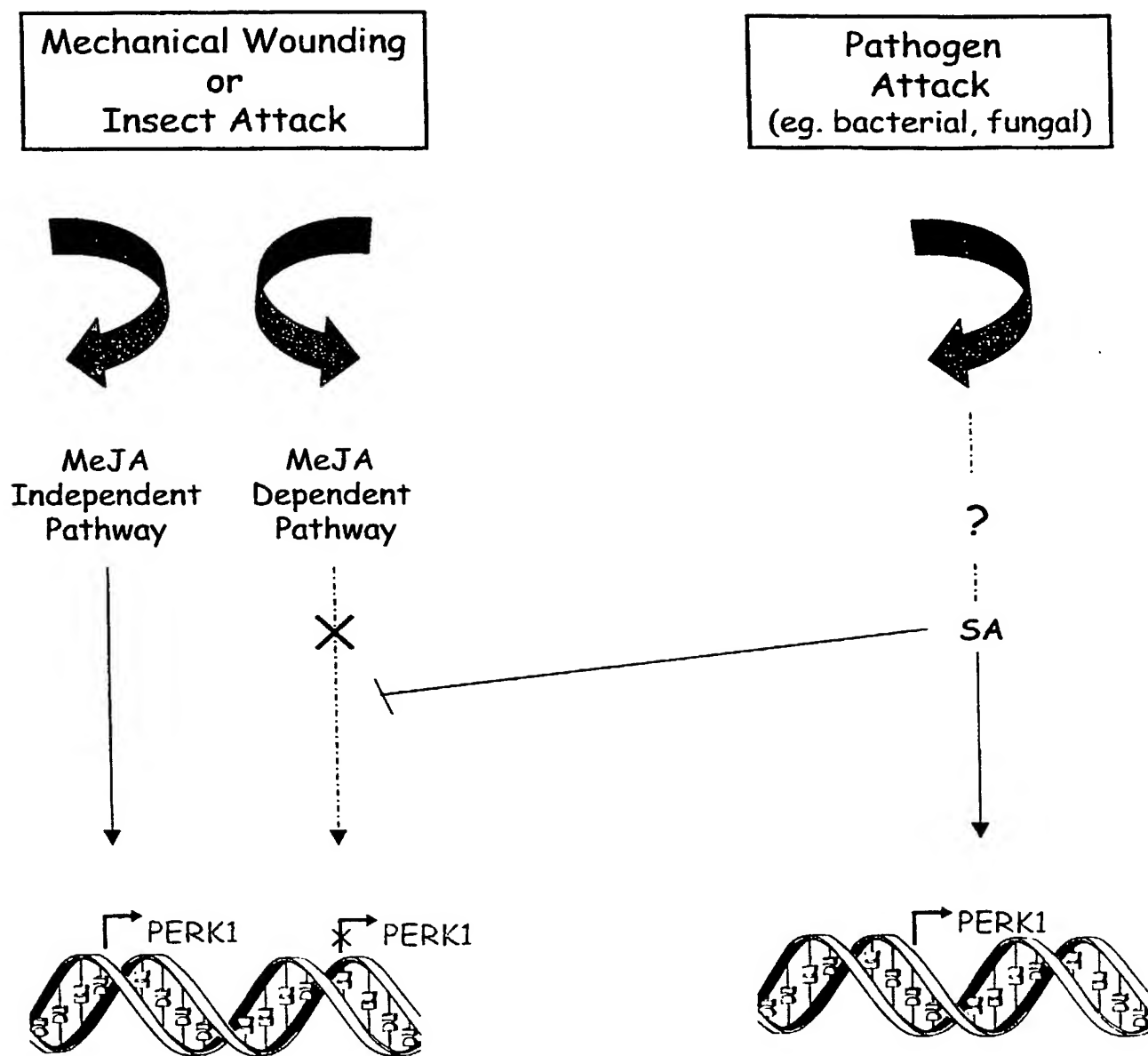


Figure 8
16/26

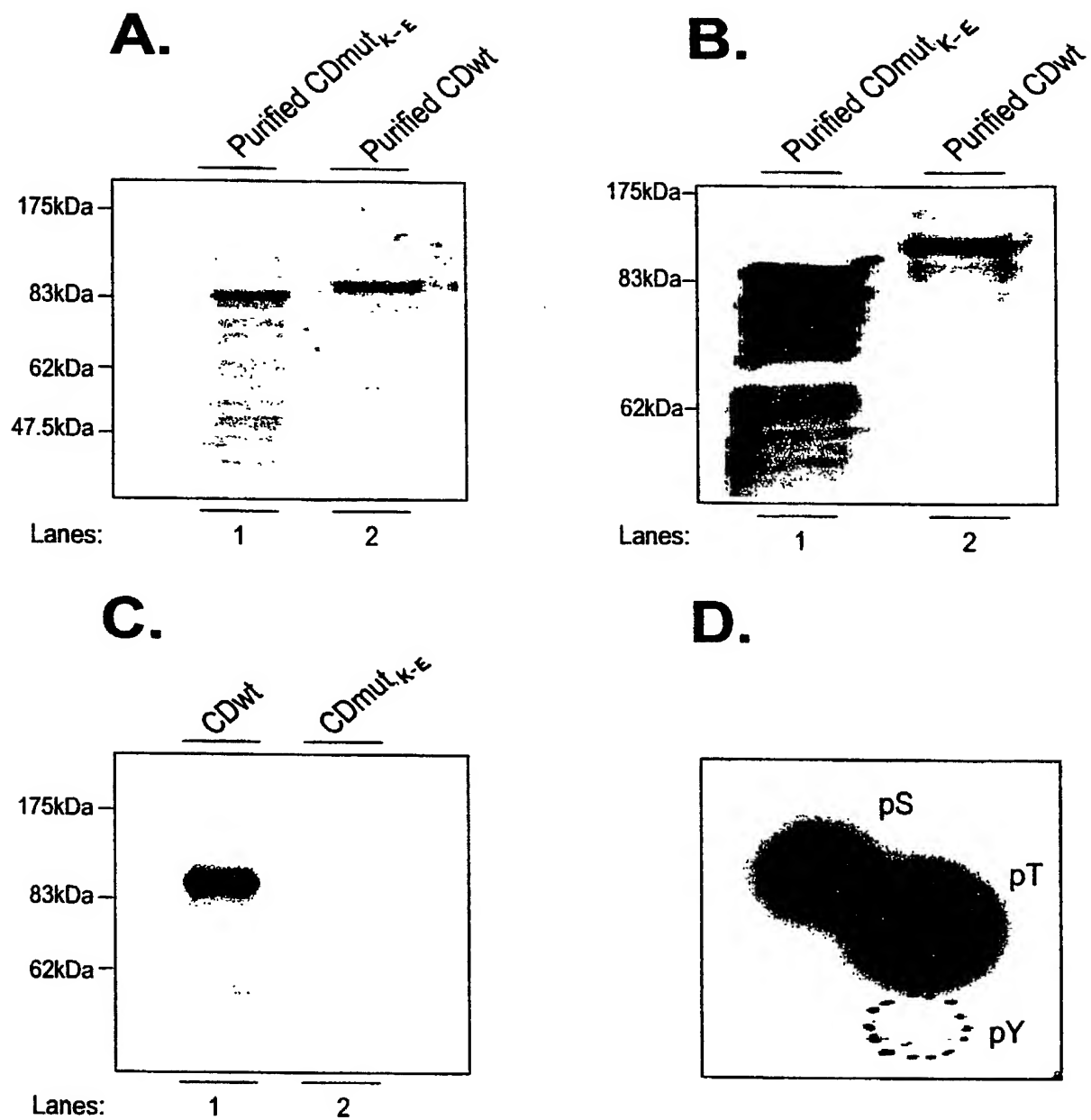
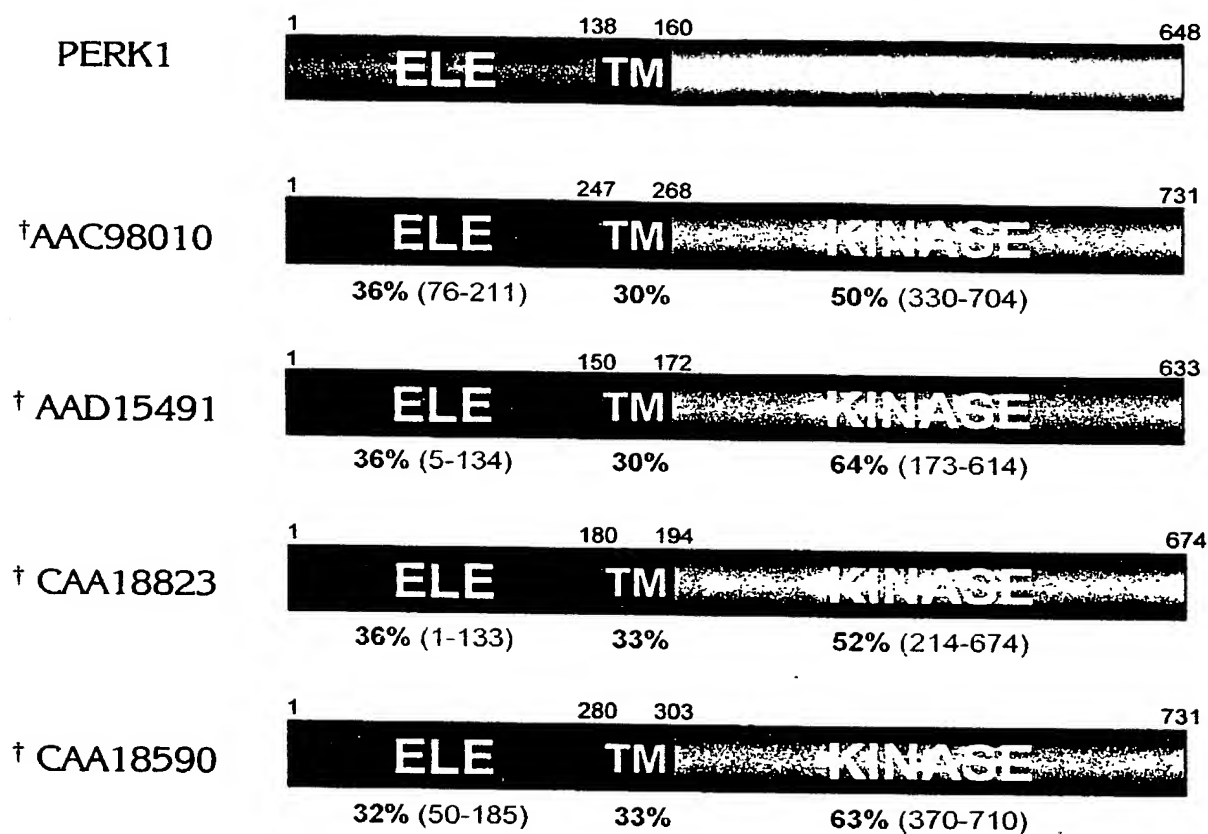


Figure 9
17/26



All have no signal peptide and predicted to be Type1b integral membrane proteins

† Predicted proteins from the Arabidopsis genome sequencing project.

ELE: Extensin-like extracellular ; TM: Transmembrane domain ; Red = sequence identity to PERK1 domain

	PERK1	ACC98010	AAD15491	CAA18823	CAA18590	
PERK1		50%	64%	52%	63%	Kinase Domains
ACC98010	36%		54%	50%	52%	
AAD15491	36%	25%		63%	51%	
CAA18823	36%	29%	27%		60%	
CAA18590	32%	36%	30%	27%		
ELE Domains						

Figure 10
18/26

TAGAAAAAAA AAAATGTCAG ACTTAGGCGA GTCGCCGAGT TCCTCACCAC CAGCACCACC
 AGCTGATACC GCTCCTCCAC CAGAGACTCC ATCAGAAAAC TCAGCTCTTC CACCTGTTGA
 TTCCTCTCCT CAGTAGCCAC CAGCTGATTC ATCATCAACA CCGCCGCTGT CAGAACCATC
 CACTCCTCCT CCAGATTCAC AGCTTCCTCC TTTACCTTCG ATTCTTCCTC CGCTAACAGA
 TTCTCCACCT CCACCTTCG ATTCTTCCTC ACCCGTTGAT TCAACCCCTT CTCCGCCGCC
 ACCGACGTC AAGCAATCTC CTCTCCTCC AGAAGATTCC GAAACACCAC CTGCTCCACC
 AAATGAATCC AATGACAACA ACCCTCCTCC GTCTCAAGAT CTTCAATCGC CTCCTCCATC
 GTCGCCGTCG CCGAATGTAG GACCCACAAA CCCGGAATCA CCACCGTTAC AATCTCCTCC
 AGCTCCACCA GCATCAGATC CTACAAATTC ACCGCCAGCT TCACCATTAG ACCCTACCAA
 TCCTCCCCCA ATACAACCAT CAGGACCAGC CACTTCTCCT CCGGCTAATC CCAACGCTCC
 GCCGAGCCCA TTCCCCACAG TACCACCCAA AACTCCTTCT AGTGGACCTG TGGTGTCTCC
 ATCTCTCACA TCCCCTAGTA AAGGAACTCC TACTCCAAAC CAAGGCAATG GAGATGGCGG
 TGGCGGTGGT GCGGCTATC AAGGGAAGAC TATGGTTGGT ATGGCTGTAG CCGGTTTCGC
 AATCATGGCG CTTATAGGCG TTGTGTTCTT AGTGAGAAGA AAGAAAAAGA GAAACATTGA
 TAGCTATAAT CACTCACAGT ACTTGCCACA TCCCAATTC TCTGTTAAAT CAGGTTTAA
 AATCTCACCT TTATCTCTCT CTGATCATCT TCTATGTGCT TGAATCATCT CTCTGACTAT
 CTTTGCTTTT GATGTAGATG GATTCTTATA CGGTCAAGAT CCAGGTAAAG GATACTCCTC
 TGGTCTAAT GGTCAATGT ATAACAATTC ACAGCAACAA CAATCCTCTA TGGGAAACAG
 TTATGGTACA GCTGGTGGTG GTTATCTCA TCATCAAATG CAATCAAGTG GCACACCTGA
 CTCTGCTATA CTCGSAAGTG GCCAGACTCA TTTCAAGTAC GAAGAGCTTG CTGAGATAAC
 ACAAGGCTTT GCTCGCAAAA ACATTCTTG AGAAGGCGGA TTTGGATGTG TCTATAAAGG
 TACATTGCAG GATGGTAAAG TTGTGCGGT TAAGCAGCTT AAAGCTGGAA GTGGACAAGG
 TGACCGTGAA TTCAAAGCAG AGGTTGAGAT CATCAGCCGC GTTCATCATC GCCATTGGT
 CTCTCTGGTT GGTACTGCA TTTAGACCA GCATAGATTG CTTATCTATG AGTATGTTTC
 TAATCAAACC TTGGAGCATC ATTTGCATGG TGAGTGACTT GTTACCATT TCGTTATAGA
 TAAGACTTTT TTTAGCTTT ACGTGTTAGA CTGACTCGCT TTACGCTTA GGAAGGGGT
 TGCCAGTTTT AGAGTGGTCT AAGAGAGTCC GGATCGCTAT AGGATCAGCC AAAGGGTTGG
 CATATCTTCA CGAAGACTGT AAGTAATGCC TTCACATTT CTTAGTTGTG TGCTTTGGT
 ATGCACTTCA TAGTTTAAAC AGAAGCCAAA AATCATATCC TTGTTTTATT TTACAGSTCA
 TCCGAAATC ATTCACAGAG ATATAAAGTC AGCAAATATT CTCTAGATG ATGAATATGA
 AGCTCAGGCA ATAATGAAAT CCTCCTTTT GTTAAATCTA TCTTATGACT GTAAAGTTT
 AGTTAATGAG ACTTGTTCTG TTTTGGAT GTTAGGTTG CTGATTTTG ACTTGCTAGA
 CTCAATGATA CAACACAAAC TCATGTTTCA ACTCGGGTTA TGGGAACCTT CGGGTAAGCA
 AACATTCATC ACAAATCTA CTCCAAACT GGACCTTATT GATCCAATGC CTGATGAAAA
 GTTTGTTATA TATGGCTTGA GGCAACAAAT TGGATCAAAC CTGAATCTT ATTGATCGTA
 TGGCTGCATG ACATGTTTTG TGTTAAGGTA CTTAGCGCCG GAATATGCAT CAAGTGGAAA
 ATTGACTGAT AGATCCGATG TATTCTCATT CGGGGTTGTT CTCTTAGAGC TGTAAGTGG
 ACCGAAACCA GTTGACCAGA CTCAGCCTCT AGGAGAAGAG AGTTTGGTTG AATGGGTAAG
 AATCCAATC TCAAACATTC TTCAATAATA GTAAGATTGG CCCTAGTATA CTTATATAGT
 ACTTATAAAT GAATCACAG GCGCGCCCGC TGCTTCTCAA AGCCATTGAG ACCGGAGATT
 TAAGCGAACT GATTGATACA CGGCTTGAAA AGCGTTATGT GGAGCATGAA GTCTTCAGAA
 TGATCGAGAC AGCCGCTGCA TGTGTTAGAC ATTCTGGTCC AAAACGCCCA CGCATGGTTC
 AGGTAATTCT GACTAACCA AAGTCCAAAG CTCCATATA TAGTAACAAG TGATTTCTCA
 CATCTGAAAA CTTATCTACT CTTCGAAATA AGGTTGTGAG AGCATTGGAC TGCGACGGAG
 ACTCGGGAGA TATTAGCAAC GGAATCAAAA TTGGGCAAAG CACAATTAT GACTCAGGGC
 AATACAATGA AGACATTATG AAATTCAGGA AAATGGCGTT TGGTGGTGAT AACAGCGTAG
 AGTCAGGATT GTACAGTGA AACTACTCTG CCAAAAGCTC TTCAGATTTC TCAGGGAATG
 AATCTGAGAC TCGGCCTTTC AACAACCGAC GGTCTGATC ATACAATAGG TGAAAGTAAC

Figure 11A

19/26

1 M S D L G E S P S S S P P A P P A D T A 20
1 ATG TCA GAC TTA GGC GAG TCG CCG AGT TCT TCA CCA CCA GCA CCA CCA GCT GAT ACC GCT 60
21 P P P E T P S E N S A L P P V D S S S P P 40
61 CCT CCA CCA GAG ACT CCA TCA GAA AAC TCA GCT CTT CCA CCT GTT GAT TCC TCT CCT CCT 120
41 S P P A D S S S T P P L S E P S T P P P 60
121 AGT CCA CCA GCT GAT TCA TCA ACA CCG CCG CTG TCA GAA CCA TCC ACT CCT CCT CCA 180
61 D S Q L P P L P S I L P P L T D S P P P 80
181 GAT TCA CAG CTT CCT CTT TTA CCT TCG ATT CTT CCT CCG CTA ACA GAT TCT CCA CCT CCA 240
81 P S D S S P P V D S T P S P P P P T S N 100
241 CCT TCC GAT TCT TCT CCA CCC GTT GAT TCA ACC CCT TCT CCG CCG CCA CCG AGC TCA AAC 300
101 E S P S P P E D S E T P P A P P N E S N 120
301 GAA TCT CCT TCT CTT CCA GAA GAT TCC GAA ACA CCA CCT GCT CCA CCA AAT GAA TCC AAT 360
121 D N N P P P S Q D L Q S P P P S S P S P 140
361 GAC AAC AAC CCT CTT CCG TCT CAA GAT CTT CAA TCG CCT CTT CCA TCG TCG CCG TCG CCG 420
141 N V G P T N P E S P P L Q S P P A P P A 160
421 AAT GTA GGA CCC ACA AAC CCG GAA TCA CCA CCG TTA CAA TCT CCT CCA GCT CCA CCA GCA 480
161 S D P T N S P A S P L D P T N P P P I 180
481 TCA GAT CCT ACA AAT TCA CCG CCA GCT TCA CCA TTA GAC CCT ACC AAT CCT CCC CCA ATA 540
181 Q P S G P A T S P P A N P N A P P S P P 200
541 CAA CCA TCA GGA CCA GCC ACT TCT CCT CCG GAT AAT CCC AAC GCT CCG CCG AGC CCA TTC 600
201 P T V P P K T P S S G P V V S P S L T S 220
601 CCC ACA GTA CCA CCC AAA ACT CTT TCT AGT GGA CCT GTG GTG TCT CCA TCT CTC ACA TCC 660
221 P S K G T P T P N Q G N G D G G G G G 240
661 CCT AGT AAA GGA ACT CCT ACT CCA AAC CAA GGC AAT GGA GAT GGC GGT GGC GGT GGC 720
241 G Y Q G K T M V G M A V A G F A I M A L 260
721 GGC TAT CAA GGG AAG ACT ATG GTT GGT ATG GCT GTA GCC GGT TTC GCA ATC ATG CCG CTT 780
261 I G V V F L V R R K K K R N I D S Y N H 280
781 ATA GGC GTT CTG TTC TTA GTG AGA AGA AAG AAA AAG AGA AAC ATT GAT AGC TAT AAT CAC 840
281 S Q Y L P H P N F S V K S D G F L Y G Q 300
841 TCA CAG TAC TTG CCA CAT CCC AAT TTC TCT GTT AAA TCA GAT GGA TTC TTA TAC GGT CAA 900
301 D P G K G Y S S G P N G S M Y N N S Q Q 320
901 GAT CCA GGT AAA GGA TAC TCC TCT GGT CTT AAT GGT TCA ATG TAT AAC AAT TCA CAG CAA 960
321 Q Q S S M G N S Y G T A G G G Y P H H Q 340
961 CAA CAA TCC TCT ATG GGA AAC AGT TAT GGT ACA GCT GGT GGT TAT CCT CAT CAT CAA 1020
341 M Q S S G T P D S A I L G S G Q T H F S 360
1021 ATG CAA TCA AGT GGC ACA CCT GAC TCT GCT ATA CTC GGA AGT GGC CAG ACT CAT TTC AGT 1080
361 Y E E L A E I T Q G F A R K N I L G E G 380
1081 TAC GAA GAG CTT GCT GAG ATA ACA CAA GGC TTT GCT CGC AAA AAC ATT CTT GGA GAA GGC 1140
381 G F G C V Y K G T L Q D G K V V A V K Q 400
1141 GGA TTT GGA TGT GTC TAT AAA GGT ACA TTG CAG GAT GGT AAA GTT GTT GCG GTT AAG CAG 1200
401 L K A G S G Q G D R E F K A E V E I I S 420
1201 CTT AAA GCT GGA AGT GGA CAA GGT GAC CGT GAA TTC AAA GCA GAG GTT GAG ATC ATC AGC 1260
421 R V H H R L V S L V G Y C I S D G H R 440
1261 CGC GTT CAT CAT CGC CAT TTG GTC TCT CTG GTT GGT TAC TGC ATT TCA GAC CAG CAT AGA 1320
441 L L I Y E Y V S N Q T L E H H L H E W S 460
1321 TTG CTT ATC TAT GAG TAT GTT TCT AAT CAA ACC TTG GAG CAT CAT TTG CAT GAG TGG TCT 1380
461 K R V R I A I G S A K G L A Y L H E D C 480
1381 AAG AGA GTC CCG ATC GCT ATA GGA TCA GCC AAA GGG TTG GCA TAT CTT CAC GAA GAC TGT 1440
481 H P K I I H R D I K S A N I L L D D E Y 500
1441 CAT CCG AAA ATC ATT CAC AGA GAT ATA AAG TCA GCA AAT ATT CTT CTA GAT GAT GAA TAT 1500
501 E A Q A I M K S S F S L N L S Y D C K V 520
1501 GAA GCT CAG GCA ATA ATG AAA TCC TCC TTT TCG TTA AAT CTA TCT TAT GAC TGT AAA GTT 1560
521 L V A D F G L A R L N D T T Q T H V S T 540
1561 TTA GTT GCT GAT TTT GGA CTT GCT ACA CTC AAT GAT ACA ACA CAA ACT CAT GTT TCA ACT 1620
541 R V M G T F G Y L A P E Y A S S G K L T 560
1621 CGG GTT ATG GGA ACC TTC GGG TAC CTA GCG CCG GAA TAT GCA TCA AGT GGA AAA TTG ACT 1680
561 D R S D V F S P G V V L L E L V T G R K 580
1681 GAT AGA TCC GAT GTA TTC TCA TTC GGG GTT GTT CTC TTA GAG CTT GTA ACT GGA CCG AAA 1740
581 P V D Q T Q P L G E E S L V E W A R P L 600
1741 CCA GTT GAC CAG ACT CAG CCT CTA GGA GAA GAG AGT TTG GTT GAA TGG GCG CCG CCG CTG 1800
601 L L K A I E T G D L S E L I D T R L E K 620
1801 CTT CTC AAA GCC ATT GAG ACC GGA GAT TTA AGC GAA CTG ATT GAT ACA CCG CTT GAA AAG 1860
621 R Y V E H E V P R M I E T A A A C V R H 640
1861 CGT TAT GTG GAG CAT GAA GTC TTC AGA ATG ATC GAG ACA GCC GCT GCA TGT GTT AGA CAT 1920
641 S G P K R P R M V Q V V R A L D C D G D 660
1921 TCT GTT CCA AAA CCG CCA CGC ATG GTT CAG GTT GTG AGA GCA TTG GAC TGC GAC GGA GAC 1980
661 S G D I S N G I K I G Q S T T Y D S G Q 680
1981 TCG GGA GAT ATT AGC AAC GGA ATC AAA ATT GGG CAA AGC ACA ACT TAT GAC TCA GGG CAA 2040
681 Y N E D I M K F R K M A P G G D N S V E 700
2041 TAC AAT GAA GAC ATT ATG AAA TTC AGG AAA ATG CCG TTT GGT GGT GAT AAC AGC GTA GAG 2100
701 S G L Y S G N Y S A K S S S D F S G N E 720
2101 TCA GGA TTG TAC AGT AAC TAC TCT GCC AAA AGC TCT TCA GAT TTC TCA GGG AAT GAA 2160
721 S E T R P F N N R R P * 732
2161 TCT GAG ACT CCG CCT TTC AAC AAC CGA CCG TTC TGA 2196

Figure 11B

20/26

GAAAATTTTG ATCTCCGATG GCTTCTTCTC CTGAATCTGC TCCTCCAACA AACTCCACCT
 CTTCTCCATC TCCACCGTCT AATACCAATT CAACCACCTC TTCTCCGCCG GCTCCGTCTC
 CTCCTTCTCC TACACCTCCT CAAGGAGACT CATCATCATC GCCACCTCCT GATTCCACAT
 CTCCACCAGC TCCACAAGCT CCTAACCTC CTAATTCCTC TAATAACTCT CTTCCGCTC
 CGTCACAGGG CGGTGGAGGA GAAAGAGGAA ATGGAGGAAA CAATGGTGGC AATGATACTC
 CACCGTCACG CGGCTCTCCT CTTTCTCCTC CTTCTAGGAG TAATGGAGAT AATGGTGGTA
 GCAGATCATC GCCACCAGGA GAACTGGAG GCTCTCGCTC AGACAACCCT CTTCTAGCG
 GAGGAAGCAG TGGAGGAGGT GGAGGTGGAA GAAGTAATAC GAATACAGCG ATCATAGTTG
 GTGTATTAGT CGGAGCTGGA CTTTGTATGA TCGTTCTTAT TATGTGTGT CTTAGACGCA
 AAAAGAAGAG AAAAGACTCC TTCTACCCTG AACCCATGAA AGGTAAAAAC ATATACACAC
 TCTTATGTTT CAACAAATAA GAAGCTTAGA TTCTTTCATA AAATTCAGG AAACCAATAT
 CAATACTATG GAAACAACAA CAACAACAAT GCTTCACAGA ATTATCCGAA TTGGCACCTA
 AATTCACAAG GCCAAAACCA ACAATCTACT GGTGGTTGGG GAGGCGGTGG ACCATCACCG
 CCTCCTCCTC CGCGGATGCC TACAAGCGGA GAAGATTCTT CCATGTACTC AGGCCCATCA
 CGCCAGTTT TACCTCCTCC TTCGCCTGCT CTAGCCCTCG GATTCAACAA GAGCACTTTT
 ACTTACCAAG AGCTTGCGGC TGCAACAGGA GGGTTACGG ATGCTAACCT TTTGGGACAG
 GGAGGATTTG GGTATGTCCA TAAAGGAGTC TTGCCTAGCG GGAAAGAAGT AGCAGTTAAG
 AGTTTAAAG CGGGTAGCGG ACAAGGAGAG AGGGAGTTT AAGCTGAGGT CGATATCATT
 AGCCGTGTGC ATCATCGGTA TCTTGTTTCT TTGGTTGGAT ATTGCATAGC TGATGGACAG
 AGGATGTTGG TTTATGAGTT TGTTCTAAC AAACTTTGG AATATCATCT TCATGTTAG
 ACCACTTAAA AACTTTGAGT ACTAAGTTTA TTTTCTCTAA TCTATATATT CAAGAAAGTT
 GTAACCTTAA TTTGTTGTTG TAGGGAAAAA TCTTCCGGTA ATGGAGTTCT CCACTAGGTT
 GCGTATCGCC TTAGGTGCTG CGAAAGGACT CGCTTACCTT CACGAAGACT GTAAGTTTAA
 ACATTCACCA TTCTCATTTT CTTAACCAAG TTGCATAAAA CAGAGAAAGC TCTGTCTCTG
 ACTAGTGTTA TCTTTTGGC TGAGAAAATG GTGCAGGCCA TCCTCGGATC ATTCACCGCG
 ACATCAAGTC TGCAAATATT CTCTGGACT TCAACTTTGA TGCTATGGTG ATAAACTAGT
 AGCTTGCAAT CATCTACGGT TTTTGTAA GACTACATTG ATGACATTTT GCATTTGTTT
 ATTCAGGTGG CTGATTTTGG ATTAGCTAAG TTAACATCTG ATAACAACAC TCATGTATCT
 ACTCGTGTA TGGGAACCTT CGGGTAAGCG TTTTACCGTA TGATAAGATT GTTCGTGACA
 CTCAGAAAC ATAACCTTTG TAGACTAATC TACTTTGCTT CTTCCACAA ACATGTGTAG
 ATATCTAGCT CCAGAATATG CTTCAAGCGG TAAATTAACC GAGAAATCCG ATGTTTTCTC
 TTACGGAGTT ATGTTATTGG AACTTATAAC TGGAAAACGA CCGGTTGATA ATAGCATCAC
 CATGGACGAC ACCTTAGTAG ATTGGGTATT CATGCATGTA ACATATGTAT CGTGTATATA
 TGTTTTTCGC CTTTTTCGCG TACTAATGAT CATGAATACA GGCTCGGCTT CTTATGGCTC
 GCGCGCTAGA AGATGGAAAC TTTAATGAGC TCGCAGATGC GAGGCTTGAA GGCAACTACA
 ACCCGCAAGA AATGGCTCGA ATGGTGACTT GTGCCGCTGC TAGCATTCTG CATTGGGGC
 GTAAACGTCC AAAGATGAGC CAGGTGAATC AAAATTATAA CTAAAAGTCT ATTTTGTCA
 GAGAATAACA AACAAATGTT GTGGTTTTCA GATAGTAAGA GCGTTAGAAG GAGAAGTGTC
 CTTAGATGCT TTAACGAAG GTGTGAAGCC AGGACACAGT AACGTTTACG GGTCAATGGG
 AGCAAGCTCG GATTATAGTC AGACATCTTA CAATGCAGAC ATGAAGAAAT TCAGACAGAT
 AGCTTTGTCT AGCCAAGAAT TCCAGTCAG TGAAGTGAA GGAACATCTA GTAATGATTC
 TAGAGATATG GGAACATAA GCCCTACTCC TCCAAAATGA GATCGAATCA ATGATTCTGT

Figure 12A

21/26

i	M	A	S	S	P	E	S	A	P	P	T	N	S	T	S	S	P	S	P	P	20
1	ATG	GCT	TCT	TCT	CCT	GAA	TCT	GCT	CCT	CCA	ACA	AAC	TCC	ACC	TCT	TCT	CCA	TCT	CCA	CCG	60
21	S	N	T	N	S	T	T	S	S	P	P	A	P	S	P	P	S	P	T	P	40
61	TCT	AAT	ACC	AAT	TCA	ACC	ACC	TCT	TCT	CCG	CCG	GCT	CCG	TCT	CCT	CCT	TCT	CCT	ACA	CCT	120
41	P	Q	G	D	S	S	S	S	P	P	P	D	S	T	S	P	P	A	P	Q	60
121	CCT	CAA	GGA	GAC	TCA	TCA	TCA	TCG	CCA	CCT	CCT	GAT	TCC	ACA	TCT	CCA	CCA	GCT	CCA	CAA	180
61	A	P	N	P	P	N	S	S	N	N	S	P	S	P	P	S	Q	G	G	G	80
181	GCT	CCT	AAC	CCT	CCT	AAT	TCC	TCT	AAT	AAC	TCT	CCT	TCC	CCT	CCG	TCA	CAG	GGC	GGT	GGA	240
81	G	E	R	G	N	G	G	N	N	G	G	N	D	T	P	P	S	R	G	S	100
241	GGA	GAA	AGA	GGA	AAT	GGA	GGA	AAC	AAT	GGT	GGC	AAT	GAT	ACT	CCA	CCG	TCA	CGC	GGC	TCT	300
101	P	P	S	P	P	S	R	S	N	G	D	N	G	G	S	R	S	S	P	P	120
301	CCT	CCT	TCT	CCT	CCT	TCT	AGG	AGT	AAT	GGA	GAT	AAT	GGT	GGT	AGC	AGA	TCA	TCG	CCA	CCA	360
121	G	D	T	G	G	S	R	S	D	N	P	P	S	S	G	G	S	S	G	G	140
361	GGA	GAC	ACT	GGA	GGC	TCT	CGC	TCA	GAC	AAC	CCT	CCT	TCT	AGC	GGA	GGA	AGC	AGT	GGA	GGA	420
141	G	G	G	G	R	S	N	T	N	T	A	I	I	V	G	V	L	V	G	A	160
421	GGT	GGA	GGT	GGA	AGA	AGT	AAT	ACG	AAT	ACA	GCG	ATC	ATA	GTT	GGT	GTA	TTA	GTC	GGA	GCT	480
161	G	L	L	M	I	V	L	I	I	V	C	L	R	R	K	K	K	R	K	D	180
481	GGA	CTT	TTG	ATG	ATC	GTT	CTT	ATT	ATT	GTG	TGT	CTT	AGA	CGC	AAA	AAG	AAG	AGA	AAA	GAC	540
181	S	F	Y	P	E	P	M	K	G	N	Q	Y	Q	Y	Y	G	N	N	N	N	200
541	TCC	TTT	TAC	CCT	GAA	CCC	ATG	AAA	GGA	AAC	CAA	TAT	CAA	TAC	TAT	GGA	AAC	AAC	AAC	AAC	600
201	N	N	A	S	Q	N	Y	P	N	W	H	L	N	S	Q	G	Q	N	Q	Q	220
601	AAC	AAT	GCT	TCA	CAG	AAT	TAT	CCG	AAT	TGG	CAC	CTA	AAT	TCA	CAA	GGC	CAA	AAC	CAA	CAA	660
221	S	T	G	G	W	G	G	G	G	P	S	P	P	P	P	P	R	M	P	T	240
661	TCT	ACT	GGT	GGT	TGG	GGA	GGC	GGT	GGA	CCA	TCA	CCG	CCT	CCT	CCT	CCG	CGG	ATG	CCT	ACA	720
241	S	G	E	D	S	S	M	Y	S	G	P	S	R	P	V	L	P	P	P	S	260
721	AGC	GGA	GAA	GAT	TCT	TCC	ATG	TAC	TCA	GGC	CCA	TCA	CGC	CCA	GTT	TTA	CCT	CCT	CCT	TCG	780
261	P	A	L	A	L	G	F	N	K	S	T	F	T	Y	Q	E	L	A	A	A	280
781	CCT	GCT	CTA	GCC	CTC	GGA	TTT	AAC	AAG	AGC	ACT	TTT	ACT	TAC	CAA	GAG	CTT	GCG	GCT	GCA	840
281	T	G	G	F	T	D	A	N	L	L	G	Q	G	G	F	G	Y	V	H	K	300
841	ACA	GGA	GGG	TTT	ACG	GAT	GCT	AAC	CTT	TTG	GGA	CAG	GGA	GGA	TTT	GGG	TAT	GTC	CAT	AAA	900
301	G	V	L	P	S	G	K	E	V	A	V	K	S	L	K	A	G	S	G	Q	320
901	GGA	GTC	TTG	CCT	AGC	GGG	AAA	GAA	GTA	GCA	GTT	AAG	AGT	TTA	AAA	GCG	GGT	AGC	GGA	CAA	960
321	G	E	R	E	F	Q	A	E	V	D	I	I	S	R	V	H	H	R	Y	L	340
961	GGA	GAG	AGG	GAG	TTT	CAA	GCT	GAG	GTC	GAT	ATC	ATT	AGC	CGT	GTG	CAT	CAT	CGG	TAT	CTT	1020
341	V	S	L	V	G	Y	C	I	A	D	G	Q	R	M	L	V	Y	E	F	V	360
1021	GTT	TCT	TTG	GTT	GGA	TAT	TGC	ATA	GCT	GAT	GGA	CAG	AGG	ATG	TTG	GTT	TAT	GAG	TTT	GTT	1080
361	P	N	K	T	L	E	Y	H	L	H	G	K	N	L	P	V	M	E	F	S	380
1081	CCT	AAC	AAA	ACT	TTG	GAA	TAT	CAT	CTT	CAT	GGG	AAA	AAT	CTT	CCG	GTA	ATG	GAG	TTC	TCC	1140
381	T	R	L	R	I	A	L	G	A	A	K	G	L	A	Y	L	H	E	D	C	400
1141	ACT	AGG	TTG	CGT	ATC	GCC	TTA	GGT	GCT	GCG	AAA	GGA	CTC	GCT	TAC	CTT	CAC	GAA	GAC	TGC	1200
401	H	P	R	I	I	H	R	D	I	K	S	A	N	I	L	L	D	P	N	F	420
1201	CAT	CCT	CGG	ATC	ATT	CAC	CGC	GAC	ATC	AAG	TCT	GCA	AAT	ATT	CTC	TTG	GAC	TTC	AAC	TTT	1260
421	D	A	M	V	A	D	F	G	L	A	K	L	T	S	D	N	N	T	H	V	440
1261	GAT	GCT	ATG	GTG	GCT	GAT	TTT	GGA	TTA	GCT	AAG	TTA	ACA	TCT	GAT	AAC	AAC	ACT	CAT	GTA	1320
441	S	T	R	V	M	G	T	F	G	Y	L	A	P	E	Y	A	S	S	G	K	460
1321	TCT	ACT	CGT	GTG	ATG	GGA	ACT	TTC	GGA	TAT	CTA	GCT	CCA	GAA	TAT	GCT	TCA	AGC	GGT	AAA	1380
461	L	T	E	K	S	D	V	F	S	Y	G	V	M	L	L	E	L	I	T	G	480
1381	TTA	ACC	GAG	AAA	TCC	GAT	GTT	TTC	TCT	TAC	GGA	GTT	ATG	TTA	TTG	GAA	CTT	ATA	ACT	GGA	1440
481	K	R	P	V	D	N	S	I	T	M	D	D	T	L	V	D	W	A	R	P	500
1441	AAA	CGA	CCG	GTT	GAT	AAT	AGC	ATC	ACC	ATG	GAC	GAC	ACC	TTA	GTA	GAT	TGG	GCT	CGG	CCT	1500
501	L	M	A	R	A	L	E	D	G	N	F	N	E	L	A	D	A	R	L	E	520
1501	CTT	ATG	GCT	CGC	GCG	CTA	GAA	GAT	GGA	AAC	TTT	AAT	GAG	CTC	GCA	GAT	GCG	AGG	CTT	GAA	1560
521	G	N	Y	N	P	Q	E	M	A	R	M	V	T	C	A	A	A	S	I	R	540
1561	GGC	AAC	TAC	AAC	CCG	CAA	GAA	ATG	GCT	CGA	ATG	GTG	ACT	TGT	GCC	GCT	GCT	AGC	ATT	CGT	1620
541	H	S	G	R	K	R	P	K	M	S	Q	I	V	R	A	L	E	G	E	V	560
1621	CAT	TCG	GGG	CGT	AAA	CGT	CCA	AAG	ATG	AGC	CAG	ATA	GTA	AGA	GCG	TTA	GAA	GGA	GAA	GTG	1680
561	S	L	D	A	L	N	E	G	V	K	P	G	H	S	N	V	Y	G	S	L	580
1681	TCC	TTA	GAT	GCT	TTA	AAC	GAA	GGT	GTG	AAG	CCA	GGA	CAC	AGT	AAC	GTT	TAC	GGG	TCA	TTG	1740
581	G	A	S	S	D	Y	S	Q	T	S	Y	N	A	D	M	K	K	F	R	Q	600
1741	GGA	GCA	AGC	TCG	GAT	TAT	AGT	CAG	ACA	TCT	TAC	AAT	GCA	GAC	ATG	AAG	AAA	TTC	AGA	CAG	1800
601	I	A	L	S	S	Q	E	F	P	V	S	D	C	E	G	T	S	S	N	D	620
1801	ATA	GCT	TTG	TCG	AGC	CAA	GAA	TTC	CCA	GTC	AGT	GAC	TGT	GAA	GGA	ACA	TCT	AGT	AAT	GAT	1860
621	S	R	D	M	G	T	K	S	P	T	P	P	K	*							634
1861	TCT	AGA	GAT	ATG	GGA	ACT	AAG	AGC	CCT	ACT	CCT	CCA	AAA	TGA							1902

Figure 12B

22/26

TCCACCGTTT GAGAAACCCT AATAACAACA TTCAAAATGG CGGACTCACC GGTGGATTCA
 TCTCCTGCCC CTGAAACCTC AAATGGGACA CCACCGTCAA ACGGAACATC GCCGTCTAAT
 GAGTCATCGC CGCCAACACC ACCTTCTTCA CCACCACCAT CATCAATATC TGCTCCTCCG
 CCAGATATCT CCGCTTCTTT TTCACCGCCG CCTGCACCAC CAACGCAAGA AACGTCACCT
 CCTACATCTC CGTCTCATC GCCGCTGTT GTAGCTAATC CGTCACCGCA GACTCCAGAG
 AATCCTTCTC CACCTGCACC TGAAGGCTCA ACTCCTGTAA CGCCACCTGC ACCACCACAA
 ACACCGTCGA ACCAATCACC GGAAGACCA ACTCCTCCTT CTCCTGGTGC CAATGATGAC
 CGAAACAGAA CCAATGGCGG AAACAACAAC AGAGACGGCT CCACACCATC ACCACCGTCG
 TCAGGGAACA GAACCTCCGG TGACGGTGGC TCACCTTCAC CACCTCGGTC GATAAGCCCT
 CCTCAGAATA GTGGAGATTG AGACTCATCA TCGGTAATC ATCCACAAGC CAACATTGGA
 TTGATTATTG GAGTCCTTGT AGGAGCAGGG CTTTGTCTC TACTTGCAGT GTGTATTGTC
 ATCTGTGCA ACAGGAAGAA GAAGAAGAAA TCTCCTCAGG TCAACCACAT GCACTACTAC
 AATAACAATC CTTATGGAGG AGCACCTCA GGTAAATACA GTTTAGTATA ACTGGAATTT
 AATTTGTAGC CTAATGGTGT TTGATTAGGT TTCAGAACGA TCATAGTCTA ATGGTTTCTG
 CTAGCTCCAT ATGGCAAAG GATTAGATT ATAAGCTAAA GGAGATGTTG CATAGTGTAG
 GTAATGGTGG TTATTACAAG GGAACACCTC AAGATCATGT GGTGAATATG GCTGGTCAAG
 GAGGTGGGAA TTGGGTCCA CAGCAACCTG TGTCTGGTCC TCACAGTGAT GCTTCCAAC
 TAACCGGTG AACTGCTATA CCGTCACCTC AAGCTGCAAC TCTTGGTCAC AACCAAAGCA
 CTTTCACATA CGATGAACG TCCATTGCAA CAGAAGGTT CGCTCAGTCA AATTTGCTAG
 GACAAGGAGG ATTTGGGTAT GTTCATAAAG GAGTTCTGCC TAGTGGCAA GAAGTTCAG
 TGAAGAGTCT TAACTTGGA AGTGGACAAG GGAACGCGA GTTTCAGCA GAGGTGATA
 TCATTAGCCG TGTCCATCAT CGTCATCTCG TTTCTCTTGT TGGATATTGC ATCTCTGGTG
 GTCAAAGACT TTTGGTTTAT GAGTTTATAC CTAACAACAC TCTTGAATTT CATCTTCATG
 GTACATTCAT CTAACAGAAT GTTTCTTGT ATTAACAAA CCTTTAAGTA TGGTTTCTCT
 TTAATCAGGA ACATGATTGA AATTTAGGA AAGGGTCGTC CGGTTTTGGA TTGGCCTACA
 AGAGTGAAGA TTGCATTGGG ATCAGCTAGA GGCCTTGCAT ATTTGCATGA AGACTGTAAG
 AAAATCTTTA TCTCACATAT TTGCATCAGT TTCTATCTCG CTCTCTACAA TATTTGAAAG
 ATTGTATATG TTACATCAAT TATAGGTCAC CCTCGCATT TCCACAGAGA TATCAAAGCT
 GCAAACATTC TTCTTGATT CAGTTTTGAG ACCAAGGTAT GTGTGTATAT ATCGACTCTT
 GTACTACTTT TACTTTCATT GTCTCTCATT TTTGTTTCCA ATCTGTGTCG ATGTGTGTAT
 CAGTCTTATT GTGTAAATAT ATGCAGGTGG CAGATTTTGG ATTGGCTAAG CTATCTCAAG
 ACAACTATAC TCATGTCTCC ACTCGCGTCA TGGGAACTTT TGGGTAAGCA GCTTTGTAAA
 ATGTCTCAAC TCATCCACAC TTATTTAGTT TCTTCACTT GTTTTAAACA TTTTCTTGGA
 TTCAGATACT TAGCTCCAGA GTATGCATCA AGCGGAAAGT TATCCGACAA ATCTGATGTT
 TTCTCAATTG GAGTAATGCT TCTTGAGCTC ATAACCGGAA GACCTCCTCT GGATCTAATC
 GGAGAAATGG AAGATAGCTT GGTAGATTGG GTAAGTCGGT CCCCCTCT TCGGTTTACT
 TGTTAATCC CAAAACACTT TCCAAAGCAA AAACAGAAAC AAATCTTACT ATTGTTGTTG
 CAGGCAAGGC CTTGTGTTT GAAAGCAGCT CAAGATGGAG ATTACAACCA ATTGGCTGAT
 CCACGTCTAG AGCTAAACTA CAGTCATCAA GAGATGGTTC AAATGGCTTC TTGTGCAGCT
 GCAGCAATCA GACATTCAGC AAGAAGACGG CCTAAGATGA GCCAGGTTCA AAAACTCATA
 CCACTTGTG GTTCTATTGT TATATTTTTA CTCACAATTA ATCTTGATGA TAAATGTGAC
 ATACTAATGA ATCTTGAAAC ATGTGTATGG TAAATGAAAA GATTGTACGA GCACTAGAAG
 GAGATATGTC AATGGATGAT CTAAGTGAGG GAACAAGACC AGGACAAAGC ACGTACTTGA
 GCCCCGGGAG CGTGAGCTCA GAGTATGACG CAAGCTCGTA CACGGCAGAC ATGAAAAAGT
 TCAAGAAACT GGCGTTAGAG AATAAGAAT ATCAAAGCAG TGAATATGGT GGAACAAGTG
 AGTATGGCTT AAACCCTTCT GCTTCAAGTA GTGAAGAAAT GAATAGAGGC TCAATGAAAC
 GCAATCCTCA GCTTGAAG AAGAGACAAC ACTTGTCATA ATATTCAGT TTTCTTCTCT

Figure 13A

23/26

10/069304

1	M	A	D	S	P	V	D	S	S	P	A	P	E	T	S	N	G	T	P	P	20
1	ATG	GCG	GAC	TCA	CCG	GTG	GAT	TCA	TCT	CCC	GCC	CCT	GAA	ACC	TCA	AAT	GGG	ACA	CCA	CCG	60
21	S	N	G	T	S	P	S	H	E	S	S	P	P	T	P	P	S	S	P	P	40
61	TCA	AAC	GGA	ACA	TCG	CCG	TCT	AAT	GAG	TCA	TCG	CCG	CCA	ACA	CCA	CCT	TCT	TCA	CCA	CCA	120
41	P	S	S	I	S	A	P	P	P	D	I	S	A	S	F	S	P	P	P	A	60
121	CCA	TCA	TCA	ATA	TCT	GCT	CCT	CCG	CCA	GAT	ATC	TCC	GCT	TCT	TTT	TCA	CCG	CCG	CCT	GCA	180
61	P	P	T	Q	E	T	S	P	P	T	S	P	S	S	S	P	P	V	V	A	80
181	CCA	CCA	ACG	CAA	GAA	ACG	TCA	CCT	CCT	ACA	TCT	CCG	TCC	TCA	TCG	CCG	CCT	GTT	GTA	GCT	240
81	N	P	S	P	Q	T	P	E	N	P	S	P	P	A	P	E	G	S	T	P	100
241	AAT	CCG	TCA	CCG	CAG	ACT	CCA	GAG	AAT	CCT	TCT	CCA	CCT	GCA	CCT	GAA	GGC	TCA	ACT	CCT	300
101	V	T	P	P	A	P	P	Q	T	P	S	N	Q	S	P	E	R	P	T	P	120
301	GTA	ACG	CCA	CCT	GCA	CCA	CAA	ACA	CCG	TCG	AAC	CAA	TCA	CCG	GAA	AGA	CCA	ACT	CCT		160
121	P	S	P	G	A	N	D	R	N	R	T	N	G	G	N	N	N	R	D		140
361	CCT	TCT	CCT	GGT	GCC	AAT	GAT	GAC	CGA	AAC	AGA	ACC	AAT	GGC	GGA	AAC	AAC	AGA	GAC		420
141	G	S	T	P	S	P	P	S	S	G	N	R	T	S	G	D	G	S	P		160
421	GGC	TCC	ACA	CCA	TCA	CCA	CCG	TCG	TCA	GGG	AAC	AGA	ACT	TCC	GGT	GAC	GGT	GGC	TCA	CCT	480
161	S	P	P	R	S	I	S	P	P	Q	N	S	G	D	S	D	S	S	S	G	180
481	TCA	CCA	CCT	CCG	TCG	ATA	AGC	CCT	CCT	CAG	AAT	AGT	GGA	GAT	TCA	GAC	TCA	TCA	TCG	GGG	540
181	L	L	L	L	L	A	V	C	I	C	I	C	N	R	K	K	K	K	K		200
541	CTT	TTG	CTT	CTA	CTT	GCA	GTG	TGT	ATT	TGC	ATC	TGT	TGC	AAC	AGG	AAG	AAG	AAG	AAG	AAA	600
201	S	P	Q	V	N	H	M	H	Y	Y	N	N	N	P	Y	G	G	A	P	S	220
601	TCT	CCT	CAG	GTC	AAC	CAC	ATG	CAC	TAC	TAC	AAT	AAC	AAT	CCT	TAT	GGA	GGA	GCA	CCC	TCA	660
221	G	N	G	Y	Y	K	G	T	P	Q	D	H	V	V	N	M	A	G	Q		240
661	GGT	AAT	GGT	GGT	TAT	TAC	AAG	GGA	ACA	CCT	CAA	GAT	CAT	GTG	GTG	AAT	ATG	GCT	GGT	CAA	720
241	G	G	G	N	W	G	P	Q	Q	P	V	S	G	P	H	S	D	A	S	N	260
721	GGA	GGT	GGG	AAT	TGG	GGT	CCA	CAG	CAA	CCT	GTG	TCT	GGT	CCT	CAC	AGT	GAT	GCT	TCC	AAC	780
261	L	T	G	R	T	A	I	P	S	P	Q	A	A	T	L	G	H	N	Q	S	280
781	TTA	ACC	GGT	CGA	ACT	GCT	ATA	CCG	TCA	CCT	CAA	GCT	GCA	ACT	CTT	GGT	CAC	AAC	CAA	AGC	840
281	T	F	T	Y	D	E	L	S	I	A	T	E	G	F	A	Q	S	N	L	L	300
841	ACT	TTC	ACA	TAC	GAT	GAA	CTG	TCC	ATT	GCA	ACA	GAA	GGT	TTC	GCT	CAG	TCA	AAT	TTG	CTA	900
301	G	Q	G	G	F	G	Y	V	H	K	G	V	L	P	S	G	K	E	V	A	320
901	GGA	CAA	GGA	GGA	TTT	GGG	TAT	GTT	CAT	AAA	GGA	GTT	CTG	CCT	AGT	GGC	AAA	GAA	GTT	GCA	960
321	V	K	S	L	K	L	G	S	G	Q	G	E	R	E	F	Q	A	E	V	D	340
961	GTG	AAG	AGT	CTT	AAA	CTT	GGA	AGT	GGA	CAA	GGG	GAA	CGC	GAG	TTT	CAA	GCA	GAG	GTT	GAT	1020
341	I	I	S	R	V	H	H	R	H	L	V	S	L	V	G	Y	C	I	S	G	360
1021	ATC	ATT	AGC	CGT	GTC	CAT	CGT	CAT	CTC	GTT	TCT	CTT	GTT	GGA	TAT	TGC	ATC	TCT	GGT		1080
361	G	Q	R	L	L	V	Y	E	F	I	P	N	N	T	L	E	F	H	L	H	380
1081	GGT	CAA	AGA	CTT	TTG	GTT	TAT	GAG	TTT	ATA	CCT	AAC	AAC	ACT	CTT	GAA	TTT	CAT	CTT	CAT	1140
381	G	K	G	R	P	V	L	D	W	P	T	R	V	K	I	A	L	G	S	A	400
1141	GGA	AAG	GGT	CGT	CCG	GTT	TTG	GAT	TGG	CCT	ACA	AGA	GTG	AAG	ATT	GCA	TTG	GGA	TCA	GCT	1200
401	R	G	L	A	Y	L	H	E	D	C	K	K	I	F	I	S	H	I	C	I	420
1201	AGA	GGC	CTT	GCA	TAT	TTG	CAT	GAA	GAC	TGT	AAG	AAA	ATC	TTT	ATC	TCA	CAT	ATT	TGC	ATC	1260
421	S	H	P	R	I	I	H	R	D	I	K	A	A	N	I	L	L	D	F	S	440
1261	AGT	CAC	CCT	CGC	ATT	ATC	CAC	AGA	GAT	ATC	AAA	GCT	GCA	AAC	ATT	CTT	CTT	GAT	TTC	AGT	1320
441	F	E	T	K	V	A	D	F	G	L	A	K	L	S	Q	D	N	Y	T	H	460
1321	TTT	GAG	ACC	AAG	GTG	GCA	GAT	TTT	GGA	TTG	GCT	AGA	CTA	TCT	CAA	GAC	AAC	TAT	ACT	CAT	1380
461	V	S	T	R	V	M	G	T	F	G	Y	L	A	P	E	Y	A	S	S	G	480
1381	GTC	TCC	ACT	CGC	GTC	ATG	GGA	ACT	TTT	GGA	TAC	TTA	GCT	CCA	GAG	TAT	GCA	TCA	AGC	GGA	1440
481	K	L	S	D	K	S	D	V	F	S	F	G	V	M	L	L	E	L	I	T	500
1441	AAG	TTA	TCC	GAC	AAA	TCT	GAT	GTT	TTC	TCA	TTT	GGA	GTA	ATG	CTT	CTT	GAG	CTC	ATA	ACC	1500
501	G	R	P	P	L	D	L	T	G	E	M	E	D	S	L	V	D	W	A	R	520
1501	GGA	AGA	CCT	CCT	CTG	GAT	CTA	ACT	GGA	GAA	ATG	GAA	GAT	AGC	TTG	GTA	GAT	TGG	GCA	AGG	1560
521	P	L	C	L	K	A	A	Q	D	G	D	Y	N	Q	L	A	D	P	R	L	540
1561	CCT	TTG	TGT	TTG	AAA	GCA	GCT	CAA	GAT	GGA	GAT	TAC	AAC	CAA	TTG	GCT	GAT	CCA	CGT	CTA	1620
541	E	L	N	Y	S	H	Q	E	M	V	Q	M	A	S	C	A	A	A	A	I	560
1621	GAG	CTA	AAC	TAC	AGT	CAT	CAA	GAG	ATG	GTT	CAA	ATG	GCT	TCT	TGT	GCA	GCT	GCA	GCA	ATC	1680
561	R	H	S	A	R	R	R	P	K	M	S	Q	V	Q	K	L	I	P	L	V	580
1681	AGA	CAT	TCA	GCA	AGA	AGA	CGG	CCT	AAG	ATG	AGC	CAG	GTT	CAA	AAA	CTC	ATA	CCA	CTT	GTT	1740
581	G	S	I	I	V	R	A	L	E	G	D	M	S	M	D	L	S	E	G		600
1741	GGT	TCT	ATT	ATT	GTA	CGA	GCA	CTA	GAA	GGA	GAT	ATG	TCA	ATG	GAT	GAT	CTA	AGT	GAG	GGA	1800
601	T	R	P	G	Q	S	T	Y	L	S	P	G	S	V	S	S	E	Y	D	A	620
1801	ACA	AGA	CCA	GGA	CAA	AGC	ACG	TAC	TTG	AGC	CCC	GGG	AGC	GTG	AGC	TCA	GAG	TAT	GAC	GCA	1860
621	S	S	Y	T	A	D	M	K	K	F	K	K	L	A	L	E	N	K	E	Y	640
1861	AGC	TCG	TAC	ACG	GCA	GAC	ATG	AAA	AAG	TTC	AAG	AAA	CTG	GCG	TTA	GAG	AAT	AAA	GAA	TAT	1920
641	Q	S	S	E	Y	G	G	T	S	E	Y	G	L	N	P	S	A	S	S	S	660
1921	CAA	AGC	AGT	GAA	TAT	GGT	GGA	ACA	AGT	GAG	TAT	GGC	TTA	AAC	CCT	TCT	GCT	TCA	AGT	AGT	1980
661	E	E	M	N	R	G	S	M	K	R	N	P	Q	L							675
1981	GAA	GAA	ATG	AAT	AGA	GGC	TCA	ATG	AAA	CGC	AAT	CCT	CAG	CTT	TGA						2025

Figure 13B

24/26

SUBSTITUTE SHEET (RULE 26)

CATACATCAC AAAACGGCAT CGTTTTGATG TCGCTCTCTC CTCTTCTCTC GCGGGTCCG
 GCAACTTCCC CGCCTGCCAT GTCATTACCT CCGGCAGATT CCGTACCTGA CACGTCATCA
 CCTCCAGCTC CTCCTTTGTC TCCTCTTCCC CCACCATTGA GCTCTCTCC GCGTTGCCT
 TCACCACCGC CTCTCTCCGC TCCCACCGCT TCCCACCGC CTCTCCGGT TGAATCCCCA
 CCGTCTCTC CTATAGAATC ACCACCGCCT CCTCTACTGG AATCACCTCC TCCTCTCCG
 TTGGAATCTC CATCTCCACC GTCTCTCAC GTCTCAGCTC CTTCGGGTTT ACCGCCATTA
 CCCTTCTCTC CGGCCAAACC TTCTCCGCCG CTTCTTTCAC CTCCCTCCGA GACAGTCCG
 CCGGGAAATA CGATTTCTCC ACCACCTCGT TCACTTCCCT CCGAATCAAC CCCGCCGGTG
 AACACAGCTT CTCTCCACC GCCATCTCCT CTCGCCGCC GTAGTGGCCC TAAGCCTTCG
 TTTCTCTCTC CCATCAATTC TTCTCCACCA AATCCTTCTC CGAACACTCC GTCACTCCCA
 GAAACTTCTC CTCCACCTAA ACCACCGCTC TCAACGACGC CATTTCCTC CTCATCCACT
 CCCCCGCTA AGAAGTCCCC TGCAGCAGTA ACTCTTCTT TCTTTGGGCC AGCGGGCCAA
 TTACCGGATG GGACCGTAGC ACCTCCTATT GGGCCTGTTA TAGAGCCCAA GACGAGTCCA
 GCCGAATCAA TATCTCCGGG AACGCCACAG CCACTGGTTC CGAAGAGTCT ACCTGTAACG
 ACGTCGTATC ACCGATCATC CGCCGGATTG TTATTTGGCG GTGTAATCGT TGGAGCTCTT
 CTACTAATTC TGTAGGTCT TCTCTTTGTC TTCTACAGAG CTACCAGAAA TAGAAATAAC
 AACAGCAGCT CTGCTCATCA TCAATCCAAA ACTCCCTCAA AAGGTATAAA CTTTGAGATC
 AATTTGTTTC AGACTGTCAC TTAAAATGTA ATGTAATTGT AAGTTATGAT CTTTTTGTCT
 TGGATCAGTT CAACATCATC GGGCGGTAA TGCTGGTACG AACAGGCAC ATGTTATCAC
 AATGCCACCA CCAATCCATG CTAATATATAT ATCTAGTGA GGTGTGATA CGAAGGAGAA
 CAATTCTGTT GCGAAAAACA TTTCAATGCC ATCTGGAATG TTCTCTACG AAGAACTTTC
 AAAAGCAACT GGTGGATTTT CAGAGGAGAA CTTTGGGA GAAGGCGGT TCGGATATGT
 TCACAAAGGA GTGTTGAAAA ACGGGACAGA AGTTGCGGTG AAGCAGCTGA AGATTGGGAG
 CTATCAAGGG GAAAGAGAAT TCCAAGCTGA GGTGACACA ATCAGTAGGG TTCATCATAA
 GCACCTCGTT TCATTGGTTG GTTATIGCGT TAATGGAGAT AAAAGACTCT TGGTTACGA
 GTTGTTCCT AAAGATACCT TGGAGTTCCA CTGTCATGGT AAAATAGATA TATGATTTC
 TCCTTTTGA TTTTGTCTT AGTTCATTAT GTTGAGTATT GTGAGAATAT GTGTTGTAGA
 GAACAGAGGA AGCGTGTGG AATGGGAAAT GAGGCTCAGG ATTGCTGTAG GAGCAGCAAA
 AGGATTAGCT TATCTTCATG AGGATTGTGA GTTGTTCCTC TTCATAATGG GAATGACAAT
 GGCCTTTTTT GGGTTTTGCA CACTGATATT GATTTCTGGT GCTGTAATG CAGGAGTCC
 AACTATAATT CACCGTGATA TCAAAGCAGC TAATATCCTT CTAGATTCCA AATTGAGGC
 AAAGGTGATC TGCTCTTAA TCTTATCAAA GTTGTGTTT TAGAACAGAG TTTGCCATT
 TTCGGTTTAT AGTCACACCA TTTGTCTTA TTTAACAAGA TTACATGTAG CTTAGAGTTC
 TTGTCTCTT CAGGTCTCTG ACTTTGACT AGCCAAGTTT TTCTCAGACA CCAATTCATC
 ATTCATCAT ATCTCTACTC GAGTGGTAGG AACTTTCCGG TAAAACACCA TCCATCCATG
 CTTTATATGT TGTGTCATTG TGTTAAAAA TATTAATTTA CGGTTTCAAG TTTCAAGTTT
 CAAGTTTCAA GTTCTTCTCT TGTGTATCG TAAATCCAGA TCAAAGATT TATTGATTAC
 TAAATGCCTT GTGCACCGTC TATTTGGTAT GCATTTAAAC AGATACATGG CTCCAGAATA
 CGCGTCCAGT GGTAAAGTAA CTGATAAATC AGATGTATAT TCCTTTGGGG TCGTGCTTCT
 AGAACTCATC ACTGGACGTC CATCAATTTT CGCCAAAGAT TCTCCACAA ACCAGAGTTT
 AGTAGACTGG GTAAGTCAAA GTACATGATG ATGATGATGA TACCATTAGG TTTCTTTT
 CCCACTAGTT ATTAGATGAA AATGAACTC CAGTATAACT GCAAGCTTGA AACTTCATTA
 GACTGTAAAT TTTGATTATT CCTCAGGCGA GGCCATTGCT TACGAAAGCA ATCTCTGGAG
 AAAGTTTGA CTTCTTGTA GACTCAAGGT TGGAGAAGAA TTACGATACA ACTCAGATGG
 CAAACATGGC TGCTGTGCT GCTGCTGCA TACGCCAATC AGCTTGGCTT CGGCCTAGAA
 TGAGCCAGGT CTGAGATTGA GTTTAAATAC ATGTATTCGT CCATATATCC GAAAGGAACA
 AGCCTAATCC ATGAATACAT TTATATCTTG AAAAAAAT GAATAGGTAG TACGTGCTCT
 TGAAGGCGAG GTGGCCCTGA GAAAGGTCGA AGAGACTGGG AATAGCGTGA CCTATAGCTC
 TTCTGAAAAC CCGAATGACA TCACACCACG GTATGGAACA AATAAGAGGA GATTGACAC
 AGGTTCAAGC GATGGTTACA CTTCAGAATA TGGAGTTAAC CTTCTCAGT CGAGCAGTGA
 ACATCAACAG GTGAATACTT AGTTCACAGG TTCAATAGGG CAAGTTTCAC CACAATTATT

Figure 14A

25/26

SUBSTITUTE SHEET (RULE 26)

1	M	S	L	S	P	S	S	S	P	A	P	A	T	S	P	P	A	M	S	L	20	
1	ATG	TCC	CTC	TCT	CCT	TCT	TCC	TGG	CCG	GCT	CCG	GCA	ACT	TCC	CCG	CCT	GCC	ATG	TCA	TTA	60	
21	P	P	A	D	S	V	P	D	T	S	S	P	P	A	P	P	L	S	P	L	40	
61	CCT	CCG	GCA	GAT	TCC	GTA	CCT	GAC	ACG	TCA	TCA	CCT	CCA	GCT	CCT	CCT	TTG	TCT	CCT	CTT	120	
41	P	P	P	L	S	S	P	P	P	L	P	S	P	P	P	P	L	S	A	P	T	60
121	CCC	CCA	CCA	TTG	AGC	TCT	CCT	CCG	CCG	TTG	CCT	TCA	CCA	CCG	CCT	CTC	TCC	GCT	CCC	ACC	180	
61	A	S	P	P	P	L	P	V	E	S	P	P	S	P	P	I	E	S	P	P	80	
181	GCT	TCC	CCA	CCG	CCT	CTT	CCG	GTT	GAA	TCC	CCA	CCG	TCT	CCT	CCT	ATA	GAA	TCA	CCA	CCG	240	
81	P	P	L	L	E	S	P	P	P	P	P	L	E	S	P	S	P	P	S	P	100	
241	CCT	CCT	CTA	CTG	GAA	TCA	CCT	CCT	CCT	CCG	TTG	GAA	TCT	CCA	TCT	CCA	CCG	TCT	CCT	CCT	300	
101	H	V	S	A	P	S	G	S	P	P	L	P	F	L	P	A	K	P	S	P	120	
301	CAC	GTC	TCA	GCT	CCT	TCC	GGT	TCA	CCG	CCA	TTA	CCC	TTT	CTT	CCC	GCC	AAA	CCT	TCT	CCG	360	
121	P	P	S	S	P	P	S	E	T	V	P	P	G	N	T	I	S	P	P	P	140	
361	CCG	CCT	TCT	TCA	CCT	CCC	TCC	GAG	ACA	GTT	CCG	CCG	GGA	AAT	ACG	ATT	TCT	CCA	CCA	CCT	420	
141	R	S	L	P	S	E	S	T	P	P	V	N	T	A	S	P	P	P	P	S	160	
421	CGT	TCA	CTT	CCC	TCC	GAA	TCA	ACC	CCG	CCG	GTG	AAC	ACA	GCT	TCT	CCT	CCA	CCG	CCA	TCT	480	
161	P	P	R	R	R	R	S	G	P	K	P	S	P	P	P	P	I	N	S	S	180	
481	CCT	CCT	CGC	CGC	CGT	AGT	GGC	CCT	AAG	CCT	TGG	TTT	CCT	CCT	CCC	ATC	AAT	TCT	TCT	CCA	540	
181	P	N	P	S	P	N	T	P	S	L	P	E	T	S	P	P	P	K	P	P	200	
541	CCA	AAT	CCT	TCT	CCG	AAC	ACT	CCG	TCA	CTC	CCA	GAA	ACT	TCT	CCT	CCA	CCT	AAA	CCA	CCG	600	
201	L	S	T	T	P	F	P	S	S	S	T	P	P	P	P	K	K	S	P	A	220	
601	CTC	TCA	ACG	ACG	CCA	TTT	CCC	TCC	TCA	TCC	ACT	CCC	CCG	CCT	AAG	AAG	TCC	CCT	GCA	GCA	660	
221	V	T	L	P	F	F	G	P	A	G	Q	L	P	D	G	T	V	A	P	P	240	
661	GTA	ACT	CTT	CCT	TTT	TTT	GGG	CCA	GCG	GGC	CAA	TTA	CCG	GAT	GGG	ACC	GTA	GCA	CCT	CCT	720	
241	I	G	P	V	I	E	P	K	T	S	P	A	E	S	I	S	P	G	T	P	260	
721	ATT	GGG	CCT	GTT	ATA	GAG	CCC	AAG	ACG	AGT	CCA	GCC	GAA	TCA	ATA	TCT	CCG	GGA	ACG	CCA	780	
261	Q	P	L	V	P	K	S	L	P	V	T	T	S	Y	H	R	S	S	A	G	280	
781	CAG	CCA	CTG	GTT	CCG	AAG	AGT	CTA	CCT	GTA	ACG	ACG	TGC	TAT	CAC	CGA	TCA	TCC	GCC	GGA	840	
281	F	L	F	G	G	V	I	V	G	A	L	L	L	I	L	L	G	L	L	F	300	
841	TTT	TTA	TTT	GGC	GGT	GTA	ATC	GTT	GGA	GCT	CTT	CTA	CTA	ATT	CTG	TTA	GGT	CTT	CTC	TTT	900	
301	V	F	Y	R	A	T	R	N	R	N	N	N	S	S	S	A	H	H	Q	S	320	
901	GTC	TTT	TAC	AGA	GCT	ACC	AGA	AAT	AGA	AAT	AAC	AAC	AGC	AGC	TCT	GCT	CAT	CAT	CAA	TCC	960	
321	K	T	P	S	K	V	Q	H	H	R	G	G	N	A	G	T	N	Q	A	H	340	
961	AAA	ACT	CCC	TCA	AAA	GTT	CAA	CAT	CAT	CCG	GGC	GGT	AAT	GCT	GGT	ACG	AAC	GAC	GCA	CAT	1020	
341	V	I	T	M	P	P	P	I	H	A	K	Y	I	S	S	G	G	C	D	T	360	
1021	GTT	ATC	ACA	ATG	CCA	CCA	ATC	CAT	GCT	AAA	TAT	ATA	TCT	AGT	GGA	GGT	TGT	GAT	ACG	1080		
361	K	E	N	N	S	V	A	K	N	I	S	M	P	S	G	M	F	S	Y	E	380	
1081	AAG	GAG	AAC	AAT	TCT	GTT	GCG	AAA	AAC	ATT	TCA	ATG	CCA	TCT	GGA	ATG	TTC	TCC	TAC	GAA	1140	
381	E	L	S	K	A	T	G	G	F	S	E	E	N	L	L	G	E	G	G	F	400	
1141	GAA	CTT	TCA	AAA	GCA	ACT	GGT	GGA	TTT	TCA	GAG	GAG	AAC	CTT	TTG	GGA	GAA	GGC	GGT	TTT	1200	
401	G	Y	V	H	K	G	V	L	K	N	G	T	E	V	A	V	K	Q	L	K	420	
1201	GGA	TAT	GTT	CAC	AAA	GGA	GTG	TTG	AAA	AAC	GGG	ACA	GAA	GTT	GCG	GTG	AAG	CAG	CTG	AAG	1260	
421	I	G	S	Y	Q	G	E	R	E	F	Q	A	E	V	D	T	I	S	R	V	440	
1261	ATT	GGG	AGC	TAT	CAA	GGG	GAA	AGA	GAA	TTT	CAA	GCT	GAG	GTT	GAC	ACA	ATC	AGT	AGG	GTT	1320	
441	H	H	K	H	L	V	S	L	V	G	Y	C	V	N	G	D	K	R	L	L	460	
1321	CAT	CAT	AAG	CAC	CTC	GTT	TCA	TTG	GTT	GCT	TAT	TGC	GTT	AAT	GGA	GAT	AAA	AGA	CTC	TTG	1380	
461	V	Y	E	P	V	P	K	D	T	L	E	P	H	L	H	E	N	R	G	S	480	
1381	GTT	TAC	GAG	TTT	GTT	CCT	AAA	GAT	ACC	TTG	GAG	TTT	CAC	TTG	CAT	GAG	AAC	AGA	GGA	AGC	1440	
481	V	L	E	W	E	M	R	L	R	I	A	V	G	A	A	K	G	L	A	Y	500	
1441	GTG	TTG	GAA	TGG	GAA	ATG	AGG	CTC	AGG	ATT	GCT	GTA	GGA	GCA	GCA	AAA	GGA	TTA	GCT	TAT	1500	
501	L	H	E	D	C	S	P	T	I	I	H	R	D	I	K	A	A	N	I	L	520	
1501	CTT	CAT	GAG	GAT	TGC	AGT	CCA	ACT	ATA	ATT	CAC	GCT	GAT	ATC	AAA	GCA	GCT	AAT	ATC	CTT	1560	
521	L	D	S	K	F	E	A	K	V	S	D	F	G	L	A	K	F	F	S	D	540	
1561	CTA	GAT	TCC	AAA	TTT	GAG	GCA	AAG	GTC	TCT	GAC	TTT	GGA	CTA	GCC	AAG	TTT	TTT	TCA	GAC	1620	
541	T	N	S	S	F	T	H	I	S	T	R	V	V	G	T	F	G	Y	M	A	560	
1621	ACC	AAT	TCA	TCA	TTT	ACT	CAT	ATC	TCT	ACT	CGA	GTG	GTA	GGA	ACT	TTT	GGA	TAC	ATG	GCT	1680	
561	P	E	Y	A	S	S	G	K	V	T	D	K	S	D	V	Y	S	F	G	V	580	
1681	CCA	GAA	TAC	GCG	TCC	AGT	GGT	AAA	GTA	ACT	GAT	AAA	TCA	GAT	GTA	TAT	TCC	TTT	GGG	GTC	1740	
581	V	L	L	E	L	I	T	G	R	P	S	I	P	A	K	D	S	S	T	N	600	
1741	GTG	CTT	CTA	GAA	CTC	ATC	ACT	GGA	CGT	CCA	TCA	ATT	TTT	GCC	AAA	GAT	TCT	TCC	ACA	AAC	1800	
601	Q	S	L	V	D	W	A	R	P	L	L	T	K	A	I	S	G	E	S	F	620	
1801	CAG	AGT	TTA	GTA	GAC	TGG	GCG	AGG	CCA	TTG	CTT	ACG	AAA	GCA	ATC	TCT	GGA	GAA	AGT	TTT	1860	
621	D	F	L	V	D	S	R	L	E	K	N	Y	D	T	T	Q	M	A	N	M	640	
1861	GAC	TTT	CTT	GTA	GAC	TCA	AGG	TTG	GAG	AAG	AAT	TAC	GAT	ACA	ACT	CAG	ATG	GCA	AAC	ATG	1920	
641	A	A	C	A	A	A	C	I	R	Q	S	A	W	L	R	P	R	M	S	Q	660	
1921	GCT	GCT	TGT	GCT	GCT	GCT	TGC	ATA	CGC	CAA	TCA	GCT	TGG	CTT	CGG	CCT	AGA	ATG	AGC	CAG	1980	
661	V	V	R	A	L	E	G	B	V	A	L	R	K	V	E	E	T	G	N	S	680	
1981	GTA	GTA	CGT	GCT	CTT	GAA	GGC	GAG	GTG	GCC	CTG	AGA	AAG	GTC	GAA	GAG	ACT	GGG	AAT	AGC	2040	
681	V	T	Y	S	S	S	E	N	P	N	D	I	T	P	R	Y	G	T	N	K	700	
2041	GTG	ACC	TAT	AGC	TCT	TCT	GAA	AAC	CCG	AAT	GAC	ATC	ACA	CCG	TAT	GGA	ACA	AAT	AAG	2100		
701	R	R	F	D	T	G	S	S	D	G	Y	T	S	E	Y	G	V	N	P	S	720	
2101	AGG	AGA	TTT	GAC	ACA	GGT	TCA	AGC	GAT	GGT	TAC	ACT	TCA	GAA	TAT	GGA	GTT	AAC	CCT	TCT	2160	
721	Q	S	S	S	E	H	Q	Q	V	N	T	*								732		
2161	CAG	TCC	AGC	AGT	GAA	CAT	CAA	CAG	GTG	AAT	ACT	TAG								2196		

Figure 14B



1111545

Docket No.
P25,762 USA

Declaration and Power of Attorney For Patent Application

English Language Declaration

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name,

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled

Proline-rich Extensin-like Receptor Kinases

the specification of which

(check one)

☐ is attached hereto.

☒ was filed on 18 August 2000 as United States Application No. or PCT International Application Number PCT/CA00/00966 and was amended on _____

(if applicable)

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, Section 119(a)-(d) or Section 365(b) of any foreign application(s) for patent or inventor's certificate, or Section 385(a) of any PCT International application which designated at least one country other than the United States, listed below and have also identified below, by checking the box, any foreign application for patent or inventor's certificate or PCT International application having a filing date before that of the application on which priority is claimed.

Prior Foreign Application(s)

Priority Not Claimed

_____ (Number)	_____ (Country)	_____ (Day/Month/Year Filed)	<input type="checkbox"/>
_____ (Number)	_____ (Country)	_____ (Day/Month/Year Filed)	<input type="checkbox"/>
_____ (Number)	_____ (Country)	_____ (Day/Month/Year Filed)	<input type="checkbox"/>

I hereby claim the benefit under 35 U.S.C. Section 119(e) of any United States provisional application(s) listed below:

60/149,466

(Application Serial No.)

60/159,122

(Application Serial No.)

(Application Serial No.)

I hereby claim the benefit under 35 U. S. C. Section 120 of any United States application(s), or Section 365(c) of any PCT International application designating the United States, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of 35 U.S.C. Section 112, I acknowledge the duty to disclose to the United States Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, C. F. R., Section 1.56 which became available between the filing date of the prior application and the national or PCT International filing date of this application:

(Application Serial No.)

(Status)
(patented, pending, abandoned)

(Application Serial No.)

(Status)
(patented, pending, abandoned)

(Application Serial No.)

(Status)
(patented; pending, abandoned)

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

POWER OF ATTORNEY: As a named inventor, I hereby appoint the following attorney(s) and/or agent(s) to prosecute this application and transact all business in the Patent and Trademark Office connected therewith. (list name and registration number)

Charles H. Lindrooth, Reg. No. 20,659

Alexis Barron, Reg. No. 22,702

Martin F. Savitzky, Reg. No. 29,699

Peter J. Butch III, Reg. No. 32,203

Joseph F. Posillico, Reg. No. 32,290

Mark D. Simpson, Reg. No. 32,942

Theodore Naccarella, Reg. No. 33,023

Patrick J. Kelly, Reg. No. 34,638

Gary A. Hecht, Reg. No. 36,826

Stephen J. Driscoll, Reg. No. 37,564

Lisa B. Lane, Reg. No. 38,217

John A. Chionchio, Reg. No. 40,254

Joseph M. Imhof, Reg. No. 41,863

Gregory S. Bernabeo, Reg. No. 44,032

Stephen J. Weed, Reg. No. 45,202

Eric H. Fischer, Reg. No. 46,010

Brett T. Freeman, Reg. No. 46,709

Gene J. Yao, Reg. No. 47,193

Jonathan M. Dermott, Reg. No. 48,608

Marc B. Bassler, Reg. No. P-48,732

Send Correspondence to: Gene J. Yao, Esquire
Synnestvedt & Lechner LLP
2600 Aramark Tower, 1101 Market Street
Philadelphia, PA 19107-2950

Direct Telephone Calls to: (name and telephone number)
Gene J. Yao - 215-923-4466

Full name of sole or first inventor

Daphne Goring

Sole or first inventor's signature

Daphne Goring

Date

April 12, 2002

Residence

2240 Taylors Orchard, Mississauga, Ontario L4E 3T3, CANADA

66 Woodhaven Crescent,
Richmond Hill, Ontario
L4E 3T2, Canada

Citizenship

Canadian

Post Office Address

Same as residence

Full name of second inventor, if any

Nancy Silva

Second inventor's signature

Nancy Silva

Date

April 12, 2002

Residence

66 Woodhaven Crescent, Richmond Hill, Ontario L4E 3T2, CANADA

2240 Taylors Orchard,
Mississauga, Ontario
L4E 3T3, Canada

Citizenship

Canadian

Post Office Address

Same as residence



PTO/PCT Rec'd 06 AUG 2002

#5

SEQUENCE LISTING

<110> GORING, Daphne R. et al.

<120> PROLINE-RICH EXTENSIN-LIKE RECEPTOR KINASES

<130> P 25,762 USA

<140> US 10/069,304

<141> 2002-02-19

<150> PCT/CA00/00966

<151> 2000-08-18

<150> US 60/149,466

<151> 1999-08-19

<150> US 60/159,122

<151> 1999-10-13

<160> 27

<170> PatentIn Ver. 2.1

<210> 1

<211> 1944

<212> DNA

<213> Brassica napus

<220>

<221> CDS

<222> (1) .. (1944)

<400> 1

atg	tcc	tcg	gcg	ccg	tct	ccg	ggg	act	ggt	tcg	cct	cca	tct	cca	cca	48
Met	Ser	Ser	Ala	Pro	Ser	Pro	Gly	Thr	Gly	Ser	Pro	Pro	Ser	Pro	Pro	
1				5					10					15		

tca	aac	tcc	aca	acc	acc	act	cct	cct	cca	gct	tcc	gct	cct	cct	ccc	96
Ser	Asn	Ser	Thr	Thr	Thr	Thr	Pro	Pro	Pro	Ala	Ser	Ala	Pro	Pro	Pro	
			20					25					30			

acc	aca	cct	tct	tct	cct	ccg	ccg	cca	tcc	act	att	ccg	aca	tct	cct	144
Thr	Thr	Pro	Ser	Ser	Pro	Pro	Pro	Pro	Ser	Thr	Ile	Pro	Thr	Ser	Pro	
		35					40					45				

cct	cct	tct	tct	cgc	tct	aca	cct	tct	gct	cct	cct	cca	tct	cca	cca	192
Pro	Pro	Ser	Ser	Arg	Ser	Thr	Pro	Ser	Ala	Pro	Pro	Pro	Ser	Pro	Pro	
	50					55					60					

act	cca	tct	acg	ccg	gga	tct	cca	cct	cct	ctt	cct	cag	ccg	tct	cca	240
Thr	Pro	Ser	Thr	Pro	Gly	Ser	Pro	Pro	Pro	Leu	Pro	Gln	Pro	Ser	Pro	
65					70					75					80	

ccc	gct	cca	act	acg	ccc	gga	tct	cca	ccc	gca	cct	gtt	act	cct	cct	288
Pro	Ala	Pro	Thr	Thr	Pro	Gly	Ser	Pro	Pro	Ala	Pro	Val	Thr	Pro	Pro	
				85					90					95		

act Thr	cga Arg	aac Asn	cct Pro 100	cca Pro	cct Pro	tca Ser	gtc Val	cca Pro 105	gga Gly	cca Pro	ccg Pro	tcc Ser	aat Asn 110	cct Pro	tca Ser	336
cgc Arg	gaa Glu	gga Gly 115	gga Gly	tct Ser	cct Pro	cga Arg	cct Pro 120	cca Pro	tct Ser	tct Ser	ccc Pro	tcg Ser 125	ccg Pro	ccg Pro	tct Ser	384
cct Pro 130	tct Ser	tcc Ser	gac Asp	ggt Gly	tta Leu	tca Ser 135	aca Thr	gga Gly	gtg Val	gtg Val	gtg Val 140	gga Gly	atc Ile	gcc Ala	atc Ile	432
gga Gly 145	gga Gly	gtc Val	gct Ala	ctg Leu	ctt Leu 150	gtg Val	ata Ile	gtg Val	act Thr	ctg Leu 155	att Ile	tgt Cys	ctc Leu	ctc Leu	tgt Cys 160	480
aag Lys	aag Lys	aaa Lys	cga Arg 165	cgg Arg	aga Arg	gac Asp	gaa Glu	gaa Glu 170	gat Asp	gct Ala	tac Tyr	tat Tyr	gtt Val	cct Pro 175	ccg Pro	528
cca Pro	cct Pro	cct Pro	cct Pro 180	ggt Gly	ccc Pro	aaa Lys	gcc Ala	gga Gly 185	gga Gly	cct Pro	tac Tyr	ggt Gly	gga Gly 190	cag Gln	cag Gln	576
caa Gln	caa Gln	tgg Trp 195	cgg Arg	caa Gln	caa Gln	aac Asn	gca Ala 200	aca Thr	cca Pro	ccg Pro	tca Ser	gat Asp 205	cat His	gtc Val	gtg Val	624
acg Thr 210	tca Ser	cta Leu	cca Pro	cca Pro	cca Pro	cct Pro 215	aag Lys	gct Ala	cca Pro	tct Ser	cca Pro 220	cca Pro	cgg Arg	caa Gln	cct Pro	672
cct Pro 225	cca Pro	cct Pro	cca Pro	cca Pro	ccg Pro 230	cct Pro	ttc Phe	atg Met	agc Ser	agc Ser 235	agc Ser	ggc Gly	ggc Gly	tcc Ser	gac Asp 240	720
tac Tyr	tcg Ser	gac Asp	cgt Arg	cca Pro 245	gtt Val	ctt Leu	cct Pro	cca Pro	ccg Pro 250	tct Ser	cca Pro	ggg Gly	ctt Leu	gtg Val 255	tta Leu	768
ggc Gly	ttc Phe	tcc Ser	aaa Lys 260	agc Ser	act Thr	ttc Phe	aca Thr	tac Tyr 265	gag Glu	gag Glu	cta Leu	gct Ala	aga Arg 270	gcc Ala	acc Thr	816
aat Asn	ggt Gly	ttc Phe 275	tcc Ser	gag Glu	gcg Ala	aac Asn	ttg Leu 280	tta Leu	gga Gly	caa Gln	ggc Gly	ggg Gly 285	ttc Phe	ggt Gly	tac Tyr	864
gtg Val 290	cac His	aaa Lys	ggt Gly	gtg Val	ttg Leu	cct Pro 295	agt Ser	ggg Gly	aaa Lys	gaa Glu	gtt Val 300	gct Ala	gtg Val	aag Lys	cag Gln	912
ttg Leu 305	aaa Lys	gtt Val	ggg Gly	agt Ser	ggt Gly 310	cag Gln	gga Gly	gag Glu	agg Arg	gag Glu 315	ttt Phe	cag Gln	gca Ala	gag Glu	gtt Val 320	960
gag	atc	atc	agc	aga	gtt	cac	cac	agg	cat	ctg	gtg	tct	ctt	gtt	ggt	1008

atg aga cca ggt caa agc aat gta tac agc tca tac gga gga agc acc	1728
Met Arg Pro Gly Gln Ser Asn Val Tyr Ser Ser Tyr Gly Gly Ser Thr	
565 570 575	
gat tat gac tcg agc cag tac aat gaa gac atg aag aag ttt agg aaa	1776
Asp Tyr Asp Ser Ser Gln Tyr Asn Glu Asp Met Lys Lys Phe Arg Lys	
580 585 590	
atg gca ctt gga act caa gag tac aac gcc acg ggt gag tac agt aat	1824
Met Ala Leu Gly Thr Gln Glu Tyr Asn Ala Thr Gly Glu Tyr Ser Asn	
595 600 605	
ccg acc agt gac tat gga ctg tac ccg tct ggt tca agc agc gag ggc	1872
Pro Thr Ser Asp Tyr Gly Leu Tyr Pro Ser Gly Ser Ser Ser Glu Gly	
610 615 620	
caa acc aca cgc gaa atg gag atg ggg aag att aag aga acc ggt cag	1920
Gln Thr Thr Arg Glu Met Glu Met Gly Lys Ile Lys Arg Thr Gly Gln	
625 630 635 640	
ggt tat agt gga cct tct ctt taa	1944
Gly Tyr Ser Gly Pro Ser Leu	
645	

<210> 2
 <211> 647
 <212> PRT
 <213> Brassica napus

<400> 2

Met Ser Ser Ala Pro Ser Pro Gly Thr Gly Ser Pro Pro Ser Pro Pro	
1 5 10 15	
Ser Asn Ser Thr Thr Thr Thr Pro Pro Pro Ala Ser Ala Pro Pro Pro	
20 25 30	
Thr Thr Pro Ser Ser Pro Pro Pro Pro Ser Thr Ile Pro Thr Ser Pro	
35 40 45	
Pro Pro Ser Ser Arg Ser Thr Pro Ser Ala Pro Pro Pro Ser Pro Pro	
50 55 60	
Thr Pro Ser Thr Pro Gly Ser Pro Pro Pro Leu Pro Gln Pro Ser Pro	
65 70 75 80	
Pro Ala Pro Thr Thr Pro Gly Ser Pro Pro Ala Pro Val Thr Pro Pro	
85 90 95	
Thr Arg Asn Pro Pro Ser Val Pro Gly Pro Pro Ser Asn Pro Ser	
100 105 110	
Arg Glu Gly Gly Ser Pro Arg Pro Ser Ser Pro Ser Pro Pro Ser	
115 120 125	
Pro Ser Ser Asp Gly Leu Ser Thr Gly Val Val Val Gly Ile Ala Ile	
130 135 140	
Gly Gly Val Ala Leu Leu Val Ile Val Thr Leu Ile Cys Leu Leu Cys	
145 150 155 160	
Lys Lys Lys Arg Arg Arg Asp Glu Glu Asp Ala Tyr Tyr Val Pro Pro	
165 170 175	
Pro Pro Pro Pro Gly Pro Lys Ala Gly Gly Pro Tyr Gly Gly Gln Gln	
180 185 190	
Gln Gln Trp Arg Gln Gln Asn Ala Thr Pro Pro Ser Asp His Val Val	
195 200 205	
Thr Ser Leu Pro Pro Pro Pro Lys Ala Pro Ser Pro Pro Arg Gln Pro	

[Illegible handwritten notes]

 $\langle 400 \rangle$ 3

tttaactctctct'gggtctccggtg	tctcctctctct	tctcctgctg	cttccttttta	acactctctt	60											
catttgccctt	tttgatttag	atccaaagaa	gcagac	atg Met 1	tcc Ser	tcg Ser	gcg Ala	ccg Pro 5	tct Ser	114						
ccg Pro	ggg Gly	act Thr	ggg Gly 10	tcg Ser	cct Pro	cca Pro	tct Ser	cca Pro 15	cca Pro	tca Ser	aac Asn	tcc Ser	aca Thr 20	acc Thr	acc Thr	162
act Thr	cct Pro	cct Pro 25	cca Pro	gct Ala	tcc Ser	gct Ala	cct Pro 30	cct Pro	ccc Pro	acc Thr	aca Thr	cct Pro 35	tct Ser	tct Ser	cct Pro	210
ccg Pro	ccg Pro 40	cca Pro	tcc Ser	act Thr	att Ile	ccg Pro 45	aca Thr	tct Ser	cct Pro	cct Pro	cct Pro 50	tct Ser	tct Ser	cgc Arg	tct Ser	258
aca Thr 55	cct Pro	tct Ser	gct Ala	cct Pro	cct Pro 60	cca Pro	tct Ser	cca Pro	cca Pro	act Thr 65	cca Pro	tct Ser	acg Thr	ccg Pro 70	gga Gly	306
tct Ser	cca Pro	cct Pro	cct Pro	ctt Leu 75	cct Pro	cag Gln	ccg Pro	tct Ser	cca Pro 80	ccc Pro	gct Ala	cca Pro	act Thr 85	acg Thr	ccc Pro	354
gga Gly	tct Ser	cca Pro	ccc Pro 90	gca Ala	cct Pro	gtt Val	act Thr	cct Pro 95	cct Pro	act Thr	cga Arg	aac Asn	cct Pro 100	cca Pro	cct Pro	402
tca Ser	gtc Val	cca Pro 105	gga Gly	cca Pro	ccg Pro	tcc Ser	aat Asn 110	cct Pro	tca Ser	cgc Arg	gaa Glu	gga Gly 115	gga Gly	tct Ser	cct Pro	450
cga Arg	cct Pro 120	cca Pro	tct Ser	tct Ser	ccc Pro	tcg Ser 125	ccg Pro	ccg Pro	tct Ser	cct Pro	tct Ser 130	tcc Ser	gac Asp	ggg Gly	tta Leu	498
tca Ser 135	aca Thr	gga Gly	gtg Val	gtg Val	gtg Val 140	gga Gly	atc Ile	gcc Ala	atc Ile	gga Gly 145	gga Gly	gtc Val	gct Ala	ctg Leu	ctt Leu 150	546
gtg Val	ata Ile	gtg Val	act Thr	ctg Leu 155	att Ile	tgt Cys	ctc Leu	ctc Leu	tgt Cys 160	aag Lys	aag Lys	aaa Lys	cga Arg	cgg Arg 165	aga Arg	594

gac Asp	gaa Glu	gaa Glu	gat Asp 170	gct Ala	tac Tyr	tat Tyr	gtt Val 175	cct Pro 175	ccg Pro	cca Pro	cct Pro	cct Pro	cct Pro	ggt Gly	ccc Pro 180	642
aaa Lys	gcc Ala	gga Gly 185	gga Gly	cct Pro	tac Tyr	ggt Gly	gga Gly 190	cag Gln	cag Gln	caa Gln	caa Gln	tgg Trp 195	cgg Arg	caa Gln	caa Gln	690
aac Asn 200	gca Ala	aca Thr	cca Pro	ccg Pro	tca Ser	gat Asp 205	cat His	gtc Val	gtg Val	acg Thr	tca Ser 210	cta Leu	cca Pro	cca Pro	cca Pro	738
cct Pro 215	aag Lys	gct Ala	cca Pro	tct Ser	cca Pro 220	cca Pro	cgg Arg	caa Gln	cct Pro	cct Pro 225	cca Pro	cct Pro	cca Pro	cca Pro	ccg Pro 230	786
cct Pro	ttc Phe	atg Met	agc Ser 235	agc Ser	agc Ser	ggc Gly	ggc Gly	tcc Ser	gac Asp 240	tac Tyr	tcg Ser	gac Asp	cgt Arg	cca Pro 245	gtt Val	834
ctt Leu	cct Pro	cca Pro	ccg Pro 250	tct Ser	cca Pro	ggg Gly	ctt Leu	gtg Val 255	tta Leu	ggc Gly	ttc Phe	tcc Ser	aaa Lys 260	agc Ser	act Thr	882
ttc Phe	aca Thr	tac Tyr 265	gag Glu	gag Glu	cta Leu	gct Ala	aga Arg 270	gcc Ala	acc Thr	aat Asn	ggt Gly	ttc Phe 275	tcc Ser	gag Glu	gcg Ala	930
aac Asn 280	ttg Leu	tta Leu	gga Gly	caa Gln	ggc Gly	ggg Gly 285	ttc Phe	ggt Gly	tac Tyr	gtg Val	cac His 290	aaa Lys	ggt Gly	gtg Val	ttg Leu	978
cct Pro 295	agt Ser	ggg Gly	aaa Lys	gaa Glu	gtt Val 300	gct Ala	gtg Val	aag Lys	cag Gln	ttg Leu 305	aaa Lys	gtt Val	ggg Gly	agt Ser	ggt Gly 310	1026
cag Gln	gga Gly	gag Glu	agg Arg	gag Glu	ttt Phe 315	cag Gln	gca Ala	gag Glu	gtt Val 320	gag Glu	atc Ile	atc Ile	agc Ser	aga Arg 325	gtt Val	1074
cac His	cac His	agg Arg	cat His 330	ctg Leu	gtg Val	tct Ser	ctt Leu	gtt Val 335	ggt Gly	tat Tyr	tgc Cys	atc Ile	gcc Ala 340	ggt Gly	gcc Ala	1122
aaa Lys	aga Arg	ttg Leu 345	ctt Leu	gtc Val	tat Tyr	gag Glu	ttt Phe 350	gtt Val	cct Pro	aac Asn	aac Asn	aat Asn 355	ctc Leu	gag Glu	ctt Leu	1170
cac His 360	ctc Leu	cat His	ggc Gly	gag Glu	gga Gly	cgg Arg 365	cct Pro	aca Thr	atg Met	gaa Glu	tgg Trp 370	agc Ser	acc Thr	aga Arg	ttg Leu	1218
aag Lys 375	att Ile	gct Ala	ctt Leu	gga Gly	tct Ser 380	gct Ala	aaa Lys	gga Gly	ctt Leu	tct Ser 385	tat Tyr	ctt Leu	cat His	gaa Glu	gat Asp 390	1266
tgc Leu	aat Ile	cct Leu	aaa Lys	atc Leu	att Leu	cac Leu	cgt Leu	gat Leu	atc Leu	aag Leu	gct Leu	tca Leu	aac Leu	ata Leu	ttg Leu	1314

THE UNIVERSITY OF CHICAGO

gag atg ggg aag att aag aga acc ggt cag ggt tat agt gga cct tct 2034
 Glu Met Gly Lys Ile Lys Arg Thr Gly Gln Gly Tyr Ser Gly Pro Ser
 635 640 645

ctt t aaaccagatg ggagagaaat tgaaggggtgt tttttcatta tttttttaaa 2088
 Leu

actgtaaaga tatgagaaaa ttgccttact ctaattaaaa ccactacgat ataaggttat 2148

aatacggtttt gaattgggttt ttaaaaaaaaa aaaaaaaaaa a 2189

<210> 4
 <211> 647
 <212> PRT
 <213> Brassica napus

<400> 4
 Met Ser Ser Ala Pro Ser Pro Gly Thr Gly Ser Pro Pro Ser Pro Pro
 1 5 10 15
 Ser Asn Ser Thr Thr Thr Thr Pro Pro Pro Ala Ser Ala Pro Pro Pro
 20 25 30
 Thr Thr Pro Ser Ser Pro Pro Pro Pro Ser Thr Ile Pro Thr Ser Pro
 35 40 45
 Pro Pro Ser Ser Arg Ser Thr Pro Ser Ala Pro Pro Pro Ser Pro Pro
 50 55 60
 Thr Pro Ser Thr Pro Gly Ser Pro Pro Pro Leu Pro Gln Pro Ser Pro
 65 70 75 80
 Pro Ala Pro Thr Thr Pro Gly Ser Pro Pro Ala Pro Val Thr Pro Pro
 85 90 95
 Thr Arg Asn Pro Pro Pro Ser Val Pro Gly Pro Pro Ser Asn Pro Ser
 100 105 110
 Arg Glu Gly Gly Ser Pro Arg Pro Pro Ser Ser Pro Ser Pro Pro Ser
 115 120 125
 Pro Ser Ser Asp Gly Leu Ser Thr Gly Val Val Val Gly Ile Ala Ile
 130 135 140
 Gly Gly Val Ala Leu Leu Val Ile Val Thr Leu Ile Cys Leu Leu Cys
 145 150 155 160
 Lys Lys Lys Arg Arg Arg Asp Glu Glu Asp Ala Tyr Tyr Val Pro Pro
 165 170 175
 Pro Pro Pro Pro Gly Pro Lys Ala Gly Gly Pro Tyr Gly Gly Gln Gln
 180 185 190
 Gln Gln Trp Arg Gln Gln Asn Ala Thr Pro Pro Ser Asp His Val Val
 195 200 205
 Thr Ser Leu Pro Pro Pro Pro Lys Ala Pro Ser Pro Pro Arg Gln Pro
 210 215 220
 Pro Pro Pro Pro Pro Pro Pro Phe Met Ser Ser Ser Gly Gly Ser Asp
 225 230 235 240
 Tyr Ser Asp Arg Pro Val Leu Pro Pro Pro Ser Pro Gly Leu Val Leu
 245 250 255
 Gly Phe Ser Lys Ser Thr Phe Thr Tyr Glu Glu Leu Ala Arg Ala Thr
 260 265 270
 Asn Gly Phe Ser Glu Ala Asn Leu Leu Gly Gln Gly Gly Phe Gly Tyr
 275 280 285
 Val His Lys Gly Val Leu Pro Ser Gly Lys Glu Val Ala Val Lys Gln
 290 295 300

His Ser Leu His Leu Pro Phe Phe Arg Ser Lys Glu Ala Asp Met Ser
20 25 30

Ser Ala Pro Ser Pro Gly Thr Gly Ser Pro Pro Ser Pro Pro Ser Asn
35 40 45

Ser Thr Thr Thr Thr Pro Pro Pro Ala Ser Ala Pro Pro Pro Thr Thr
50 55 60

Pro Ser Ser Pro Pro Pro Pro Ser Thr Ile Pro Thr Ser Pro Pro Pro
65 70 75 80

Ser Ser Arg Ser Thr Pro Ser Ala Pro Pro Pro Ser Pro Pro Thr Pro
85 90 95

Ser Thr Pro Gly Ser Pro Pro Pro Leu Pro Gln Pro Ser Pro Pro Ala
100 105 110

Pro Thr Thr Pro Gly Ser Pro Pro Ala Pro Val Thr Pro Pro Thr Arg
115 120 125

Asn Pro Pro Pro Ser Val Pro Gly Pro Pro Ser Asn Pro Ser Arg Glu
130 135 140

Gly Gly Ser Pro Arg Pro Pro Ser Ser Pro Ser Pro Pro Ser Pro Ser
145 150 155 160

Ser Asp Gly Leu Ser Thr Gly Val Val Val Gly Ile Ala Ile Gly Gly
165 170 175

Val Ala Leu Leu Val Ile Val Thr Leu Ile Cys Leu Leu Cys Lys Lys
180 185 190

Lys Arg Arg Arg Asp Glu Glu Asp Ala Tyr Tyr Val Pro Pro Pro Pro
195 200 205

Pro Pro Gly Pro Lys Ala Gly Gly Pro Tyr Gly Gly Gln Gln Gln Gln
210 215 220

Trp Arg Gln Gln Asn Ala Thr Pro Pro Ser Asp His Val Val Thr Ser
225 230 235 240

Leu Pro Pro Pro Pro Lys Ala Pro Ser Pro Pro Arg Gln Pro Pro Pro
245 250 255

[illegible]

Asp Arg Pro Val Leu Pro Pro Pro Ser Pro Gly Leu Val Leu Gly Phe
275 280 285

Ser Lys Ser Thr Phe Thr Tyr Glu Glu Leu Ala Arg Ala Thr Asn Gly
290 295 300

Phe Ser Glu Ala Asn Leu Leu Gly Gln Gly Gly Phe Gly Tyr Val His
305 310 315 320

Lys Gly Val Leu Pro Ser Gly Lys Glu Val Ala Val Lys Gln Leu Lys
325 330 335

Val Gly Ser Gly Gln Gly Glu Arg Glu Phe Gln Ala Glu Val Glu Ile
340 345 350

Ile Ser Arg Val His His Arg His Leu Val Ser Leu Val Gly Tyr Cys
355 360 365

Ile Ala Gly Ala Lys Arg Leu Leu Val Tyr Glu Phe Val Pro Asn Asn
370 375 380

Asn Leu Glu Leu His Leu His Gly Glu Gly Arg Pro Thr Met Glu Trp
385 390 395 400

Ser Thr Arg Leu Lys Ile Ala Leu Gly Ser Ala Lys Gly Leu Ser Tyr
405 410 415

Leu His Glu Asp Cys Asn Pro Lys Ile Ile His Arg Asp Ile Lys Ala
420 425 430

Ser Asn Ile Leu Ile Asp Phe Lys Phe Glu Ala Lys Val Ala Asp Phe
435 440 445

Gly Leu Ala Lys Ile Ala Ser Asp Thr Asn Thr His Val Ser Thr Arg
450 455 460

Val Met Gly Thr Phe Gly Tyr Leu Ala Pro Glu Tyr Ala Ala Ser Gly
465 470 475 480

Lys Leu Thr Glu Lys Ser Asp Val Phe Ser Phe Gly Val Val Leu Leu
485 490 495

Glu Leu Ile Thr Gly Arg Arg Pro Val Asp Ala Asn Asn Val Tyr Val
500 505 510

Asp Asp Ser Leu Val Asp Trp Ala Arg Pro Leu Leu Asn Arg Ala Ser
515 520 525

Glu Gln Gly Asp Phe Glu Gly Leu Ala Asp Ala Lys Met Asn Asn Gly
530 535 540

Tyr Asp Arg Glu Glu Met Ala Arg Met Val Ala Cys Ala Ala Ala Cys
545 550 555 560

Val Arg His Ser Ala Arg Arg Arg Pro Arg Met Ser Gln Ile Val Arg
565 570 575

Ala Leu Glu Gly Asn Val Ser Leu Ser Asp Leu Asn Glu Gly Met Arg
580 585 590

Pro Gly Gln Ser Asn Val Tyr Ser Ser Tyr Gly Gly Ser Thr Asp Tyr
595 600 605

Asp Ser Ser Gln Tyr Asn Glu Asp Met Lys Lys Phe Arg Lys Met Ala
610 615 620

Leu Gly Thr Gln Glu Tyr Asn Ala Thr Gly Glu Tyr Ser Asn Pro Thr
625 630 635 640

Ser Asp Tyr Gly Leu Tyr Pro Ser Gly Ser Ser Ser Glu Gly Gln Thr
645 650 655

Thr Arg Glu Met Glu Met Gly Lys Ile Lys Arg Thr Gly Gln Gly Tyr
660 665 670

Ser Gly Pro Ser Leu Thr Arg Trp Glu Arg Asn Arg Val Phe Phe His
675 680 685

Tyr Phe Phe Lys Thr Val Lys Ile Glu Asn Cys Leu Thr Leu Ile Lys
690 695 700

Thr Thr Thr Ile Gly Tyr Asn Thr Phe Ile Gly Phe Lys Lys Lys Lys
705 710 715 720

Lys

<210> 6

<211> 2231

<212> DNA

<213> Arabidopsis thaliana

<400> 6

tagaaaaaaaa	aaaatgtcag	acttaggcga	gtcgccgagt	tcttcaccac	cagcaccacc	60
agctgatacc	gctcctccac	cagagactcc	atcagaaaac	tcagctcttc	cacctgttga	120
ttcctctcct	cctagtcac	cagctgattc	atcatcaaca	ccgccgctgt	cagaaccatc	180
cactcctcct	ccagattcac	agcttcctcc	tttaccttcg	attcttcctc	cgctaacaga	240
ttctccacct	ccaccttcg	attcttctcc	acccgttgat	tcaaccctt	ctccgccgcc	300
accgacgtca	aacgaatctc	cttctcctcc	agaagattcc	gaaacaccac	ctgctccacc	360
aaatgaatcc	aatgacaaca	accctcctcc	gtctcaagat	cttcaatcgc	ctcctccatc	420
gtcgccgtcg	ccgaatgtag	gacccacaaa	cccggaatca	ccaccgttac	aatctcctcc	480
agctccacca	gcatcagatc	ctacaaattc	accgccagct	tcaccattag	accctaccaa	540
tcctccccc	atacaacat	caggaccagc	cacttctcct	ccggctaate	ccaacgctcc	600
gccgagccca	ttccccacag	taccacccaa	aactccttct	agtggacctg	tgggtgtctc	660
atctctcaca	tcccctagta	aaggaaactcc	tactccaaac	caaggcaatg	gagatggcgg	720
tggcgggtggt	ggcggctatc	aagggaagac	tatggttggt	atggctgtag	ccggtttcgc	780
aatcatggcg	cttataggcg	ttgtgttctt	agtgagaaga	aagaaaaaga	gaaacattga	840
tagctataat	cactcacagt	acttgccaca	tcccaatttc	tctgttaa	cagatggatt	900
cttatacggg	caagatccag	gtaaaggata	ctcctctggt	cctaattggt	caatgtataa	960
caattcacag	caacaacaat	cctctatggg	aaacagttat	ggtacagctg	gtgggtggtta	1020
tcctcatcat	caaagtgaat	caagtggcac	acctgactct	gctatactcg	gaagtggcca	1080
gactcatttc	agttacgaag	agcttgctga	gataacacaa	ggctttgctc	gcaaaaacat	1140
tcttgagaaa	ggcggatttg	gatgtgtcta	taaagggtaca	ttgcaggatg	gtaaagtgtg	1200
tgcggttaag	cagcttaaa	ctggaagtgg	acaagggtgac	cgtgaattca	aagcagaggt	1260
tgagatcatc	agccgcgttc	atcatcgcca	tttggtctct	ctggttggtt	actgcatttc	1320
agaccagcat	agattgctta	tctatgagta	tgtttctaat	caaaccttgg	agcatcattt	1380
gcatgagtgg	tctaagagag	tccggatcgc	tataggatca	gccaaagggt	tggcataatc	1440
tcacgaagac	tgtcatccga	aaatcattca	cagagatata	aagtcagcaa	atattcttct	1500
agatgatgaa	tatgaagctc	aggcaataat	gaaatcctcc	ttttcgttaa	atctatctta	1560
tgactgtaaa	gttttagttg	ctgatttttg	acttgctaga	ctcaatgata	caacacaaac	1620
tcatgtttca	actcgggtta	tgggaacctt	cgggtaccta	gcgccggaat	atgcatcaag	1680
tggaaaattg	actgatagat	ccgatgtatt	ctcattcggg	gttggttctc	tagagcttgt	1740
aactggacgg	aaaccagttg	accagactca	gcctctagga	gaagagagtt	tggttgaatg	1800
ggcgcgcccg	ctgcttctca	aagccattga	gaccggagat	ttaagcgaac	tgattgatac	1860
acggcttgaa	aagcgttatg	tggagcatga	agtcttcaga	atgategaga	cagccgctgc	1920
atgtgttaga	cattctgggtc	caaaacgccc	acgcattggt	caggttgtga	gagcaattga	1980
ctgcgacgga	gactcgggag	atattagcaa	cggaatcaaa	attgggcaaa	gcacaactta	2040
tgactcaggg	caatacaatg	aagacattat	gaaattcagg	aaaatggcgt	tgggtggtga	2100
taacagcgta	gagtcaggat	tgtacagtgg	aaactactct	gccaaaagct	cttcagattt	2160
ctcaggggaat	gaatctgaga	ctcggccttt	caacaaccga	cggttctgat	catacaatag	2220
gtgaaagtaa	c					2231

<210> 7

<211> 2196

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(2196)

<400> 7

atg	tca	gac	tta	ggc	gag	tcg	ccg	agt	tct	tca	cca	cca	gca	cca	cca	48
Met	Ser	Asp	Leu	Gly	Glu	Ser	Pro	Ser	Ser	Ser	Pro	Pro	Ala	Pro	Pro	
1				5					10					15		
gct	gat	acc	gct	cct	cca	cca	gag	act	cca	tca	gaa	aac	tca	gct	ctt	96
Ala	Asp	Thr	Ala	Pro	Pro	Pro	Glu	Thr	Pro	Ser	Glu	Asn	Ser	Ala	Leu	
			20					25					30			
cca	cct	gtt	gat	tcc	tct	cct	cct	agt	cca	cca	gct	gat	tca	tca	tca	144
Pro	Pro	Val	Asp	Ser	Ser	Pro	Pro	Ser	Pro	Pro	Ala	Asp	Ser	Ser	Ser	
		35					40					45				
aca	ccg	ccg	ctg	tca	gaa	cca	tcc	act	cct	cct	cca	gat	tca	cag	ctt	192
Thr	Pro	Pro	Leu	Ser	Glu	Pro	Ser	Thr	Pro	Pro	Pro	Asp	Ser	Gln	Leu	
	50					55					60					
cct	cct	tta	cct	tcg	att	ctt	cct	ccg	cta	aca	gat	tct	cca	cct	cca	240
Pro	Pro	Leu	Pro	Ser	Ile	Leu	Pro	Pro	Leu	Thr	Asp	Ser	Pro	Pro	Pro	
65					70					75					80	
cct	tcc	gat	tct	tct	cca	ccc	gtt	gat	tca	acc	cct	tct	ccg	ccg	cca	288
Pro	Ser	Asp	Ser	Ser	Pro	Pro	Val	Asp	Ser	Thr	Pro	Ser	Pro	Pro	Pro	
				85					90					95		
ccg	acg	tca	aac	gaa	tct	cct	tct	cct	cca	gaa	gat	tcc	gaa	aca	cca	336
Pro	Thr	Ser	Asn	Glu	Ser	Pro	Ser	Pro	Pro	Glu	Asp	Ser	Glu	Thr	Pro	
			100					105					110			
cct	gct	cca	cca	aat	gaa	tcc	aat	gac	aac	aac	cct	cct	ccg	tct	caa	384
Pro	Ala	Pro	Pro	Asn	Glu	Ser	Asn	Asp	Asn	Asn	Pro	Pro	Pro	Ser	Gln	
		115					120					125				
gat	ctt	caa	tcg	cct	cct	cca	tcg	tcg	ccg	tcg	ccg	aat	gta	gga	ccc	432
Asp	Leu	Gln	Ser	Pro	Pro	Pro	Ser	Ser	Pro	Ser	Pro	Asn	Val	Gly	Pro	
	130					135					140					
aca	aac	ccg	gaa	tca	cca	ccg	tta	caa	tct	cct	cca	gct	cca	cca	gca	480
Thr	Asn	Pro	Glu	Ser	Pro	Pro	Leu	Gln	Ser	Pro	Pro	Ala	Pro	Pro	Ala	
145					150					155					160	
tca	gat	cct	aca	aat	tca	ccg	cca	gct	tca	cca	tta	gac	cct	acc	aat	528
Ser	Asp	Pro	Thr	Asn	Ser	Pro	Pro	Ala	Ser	Pro	Leu	Asp	Pro	Thr	Asn	
				165					170					175		
cct	ccc	cca	ata	caa	cca	tca	gga	cca	gcc	act	tct	cct	ccg	gct	aat	576
Pro	Pro	Pro	Ile	Gln	Pro	Ser	Gly	Pro	Ala	Thr	Ser	Pro	Pro	Ala	Asn	
			180				185						190			
ccc	aac	gct	ccg	ccg	agc	cca	ttc	ccc	aca	gta	cca	ccc	aaa	act	cct	624
Pro	Asn	Ala	Pro	Pro	Ser	Pro	Phe	Pro	Thr	Val	Pro	Pro	Lys	Thr	Pro	
		195					200					205				
tct	agt	gga	cct	gtg	gtg	tct	cca	tct	ctc	aca	tcc	cct	agt	aaa	gga	672
Ser	Ser	Gly	Pro	Val	Val	Ser	Pro	Ser	Leu	Thr	Ser	Pro	Ser	Lys	Gly	
	210					215					220					

act Thr 225	cct Pro	act Thr	cca Pro	aac Asn	caa Gln 230	ggc Gly	aat Asn	gga Gly	gat Asp	ggc Gly 235	ggt Gly	ggc Gly	ggt Gly	ggt Gly	ggc Gly 240	720
ggc Gly	tat Tyr	caa Gln	ggg Gly	aag Lys 245	act Thr	atg Met	gtt Val	ggt Gly	atg Met 250	gct Ala	gta Val	gcc Ala	ggt Gly	ttc Phe 255	gca Ala	768
atc Ile	atg Met	gcg Ala	ctt Leu 260	ata Ile	ggc Gly	gtt Val	gtg Val	ttc Phe 265	tta Leu	gtg Val	aga Arg	aga Arg	aag Lys 270	aaa Lys	aag Lys	816
aga Arg	aac Asn	att Ile 275	gat Asp	agc Ser	tat Tyr	aat Asn	cac His 280	tca Ser	cag Gln	tac Tyr	ttg Leu	cca Pro 285	cat His	ccc Pro	aat Asn	864
ttc Phe	tct Ser 290	gtt Val	aaa Lys	tca Ser	gat Asp	gga Gly 295	ttc Phe	tta Leu	tac Tyr	ggt Gly	caa Gln 300	gat Asp	cca Pro	ggt Gly	aaa Lys	912
gga Gly 305	tac Tyr	tcc Ser	tct Ser	ggt Gly	cct Pro 310	aat Asn	ggt Gly	tca Ser	atg Met	tat Tyr 315	aac Asn	aat Asn	tca Ser	cag Gln	caa Gln 320	960
caa Gln	caa Gln	tcc Ser	tct Ser	atg Met 325	gga Gly	aac Asn	agt Ser	tat Tyr	ggt Gly 330	aca Thr	gct Ala	ggt Gly	ggt Gly	ggt Gly 335	tat Tyr	1008
cct Pro	cat His	cat His	caa Gln 340	atg Met	caa Gln	tca Ser	agt Ser	ggc Gly 345	aca Thr	cct Pro	gac Asp	tct Ser	gct Ala 350	ata Ile	ctc Leu	1056
gga Gly	agt Ser	ggc Gly 355	cag Gln	act Thr	cat His	ttc Phe	agt Ser 360	tac Tyr	gaa Glu	gag Glu	ctt Leu	gct Ala 365	gag Glu	ata Ile	aca Thr	1104
caa Gln	ggc Gly 370	ttt Phe	gct Ala	cgc Arg	aaa Lys	aac Asn 375	att Ile	ctt Leu	gga Gly	gaa Glu	ggc Gly 380	gga Gly	ttt Phe	gga Gly	tgt Cys	1152
gtc Val 385	tat Tyr	aaa Lys	ggt Gly	aca Thr	ttg Leu 390	cag Gln	gat Asp	ggt Gly	aaa Lys	gtt Val 395	gtt Val	gcg Ala	gtt Val	aag Lys	cag Gln 400	1200
ctt Leu	aaa Lys	gct Ala	gga Gly	agt Ser 405	gga Gly	caa Gln	ggt Gly	gac Asp	cgt Arg 410	gaa Glu	ttc Phe	aaa Lys	gca Ala	gag Glu 415	gtt Val	1248
gag Glu	atc Ile	atc Ile	agc Ser 420	cgc Arg	gtt Val	cat His	cat His	cgc Arg 425	cat His	ttg Leu	gtc Val	tct Ser	ctg Leu 430	gtt Val	ggt Gly	1296
tac Tyr	tgc Cys	att Ile 435	tca Ser	gac Asp	cag Gln	cat His	aga Arg 440	ttg Leu	ctt Leu	atc Ile	tat Tyr	gag Glu 445	tat Tyr	gtt Val	tct Ser	1344
aat	caa	acc	ttg	gag	cat	cat	ttg	cat	gag	tgg	tct	aag	aga	gtc	cgg	1392

Asn 450	Gln	Thr	Leu	Glu	His 455	His	Leu	His	Glu	Trp	Ser 460	Lys	Arg	Val	Arg	
atc 465	gct Ala	ata Ile	gga Gly	tca Ser	gcc Ala 470	aaa Lys	ggg Gly	ttg Leu	gca Ala 475	tat Tyr	ctt Leu	cac His	gaa Glu	gac Asp	tgt Cys 480	1440
cat His	ccg Pro	aaa Lys	atc Ile	att Ile 485	cac His	aga Arg	gat Asp	ata Ile	aag Lys 490	tca Ser	gca Ala	aat Asn	att Ile	ctt Leu 495	cta Leu	1488
gat Asp	gat Asp	gaa Glu	tat Tyr 500	gaa Glu	gct Ala	cag Gln	gca Ala	ata Ile 505	atg Met	aaa Lys	tcc Ser	tcc Ser	ttt Phe 510	tcg Ser	tta Leu	1536
aat Asn	cta Leu	tct Ser 515	tat Tyr	gac Asp	tgt Cys	aaa Lys	gtt Val 520	tta Leu	gtt Val	gct Ala	gat Asp	ttt Phe 525	gga Gly	ctt Leu	gct Ala	1584
aga Arg 530	ctc Leu	aat Asn	gat Asp	aca Thr	aca Thr	caa Gln 535	act Thr	cat His	gtt Val	tca Ser	act Thr 540	cgg Arg	gtt Val	atg Met	gga Gly	1632
acc Thr 545	ttc Phe	ggg Gly	tac Tyr	cta Leu	gcg Ala 550	ccg Pro	gaa Glu	tat Tyr	gca Ala 555	tca Ser	agt Ser	gga Gly	aaa Lys	ttg Leu	act Thr 560	1680
gat Asp	aga Arg	tcc Ser	gat Asp	gta Val 565	ttc Phe	tca Ser	ttc Phe	ggg Gly	gtt Val 570	gtt Val	ctc Leu	tta Leu	gag Glu	ctt Leu 575	gta Val	1728
act Thr	gga Gly	cgg Arg	aaa Lys 580	cca Pro	gtt Val	gac Asp	cag Gln	act Thr 585	cag Gln	cct Pro	cta Leu	gga Gly	gaa Glu 590	gag Glu	agt Ser	1776
ttg Leu	gtt Val	gaa Glu 595	tgg Trp	gcg Ala	cgc Arg	ccg Pro	ctg Leu 600	ctt Leu	ctc Leu	aaa Lys	gcc Ala 605	att Ile	gag Glu	acc Thr	gga Gly	1824
gat Asp	tta Leu 610	agc Ser	gaa Glu	ctg Leu	att Ile	gat Asp 615	aca Thr	cgg Arg	ctt Leu	gaa Glu	aag Lys 620	cgt Arg	tat Tyr	gtg Val	gag Glu	1872
cat His 625	gaa Glu	gtc Val	ttc Phe	aga Arg	atg Met 630	atc Ile	gag Glu	aca Thr	gcc Ala 635	gct Ala	gca Ala	tgt Cys	gtt Val	aga Arg	cat His 640	1920
tct Ser	ggt Gly	cca Pro	aaa Lys	cgc Arg	cca Pro	cgc Arg	atg Met	gtt Val	cag Gln 650	gtt Val	gtg Val	aga Arg	gca Ala	ttg Leu 655	gac Asp	1968
tgc Cys	gac Asp	gga Gly	gac Asp 660	tcg Ser	gga Gly	gat Asp	att Ile	agc Ser 665	aac Asn	gga Gly	atc Ile	aaa Lys	att Ile 670	ggg Gly	caa Gln	2016
agc Ser	aca Thr	act Thr 675	tat Tyr	gac Asp	tca Ser	ggg Gly	caa Gln 680	tac Tyr	aat Asn	gaa Glu	gac Asp	att Ile 685	atg Met	aaa Lys	ttc Phe	2064

agg aaa atg gcg ttt ggt ggt gat aac agc gta gag tca gga ttg tac 2112
 Arg Lys Met Ala Phe Gly Gly Asp Asn Ser Val Glu Ser Gly Leu Tyr
 690 695 700

agt gga aac tac tct gcc aaa agc tct tca gat ttc tca ggg aat gaa 2160
 Ser Gly Asn Tyr Ser Ala Lys Ser Ser Ser Asp Phe Ser Gly Asn Glu
 705 710 715 720

tct gag act cgg cct ttc aac aac cga cgg ttc tga 2196
 Ser Glu Thr Arg Pro Phe Asn Asn Arg Arg Phe
 725 730

<210> 8
 <211> 731
 <212> PRT
 <213> Arabidopsis thaliana

<400> 8
 Met Ser Asp Leu Gly Glu Ser Pro Ser Ser Ser Pro Pro Ala Pro Pro
 1 5 10 15
 Ala Asp Thr Ala Pro Pro Pro Glu Thr Pro Ser Glu Asn Ser Ala Leu
 20 25 30
 Pro Pro Val Asp Ser Ser Pro Pro Ser Pro Pro Ala Asp Ser Ser Ser
 35 40 45
 Thr Pro Pro Leu Ser Glu Pro Ser Thr Pro Pro Pro Asp Ser Gln Leu
 50 55 60
 Pro Pro Leu Pro Ser Ile Leu Pro Pro Leu Thr Asp Ser Pro Pro Pro
 65 70 75 80
 Pro Ser Asp Ser Ser Pro Pro Val Asp Ser Thr Pro Ser Pro Pro Pro
 85 90 95
 Pro Thr Ser Asn Glu Ser Pro Ser Pro Pro Glu Asp Ser Glu Thr Pro
 100 105 110
 Pro Ala Pro Pro Asn Glu Ser Asn Asp Asn Asn Pro Pro Pro Ser Gln
 115 120 125
 Asp Leu Gln Ser Pro Pro Pro Ser Ser Pro Ser Pro Asn Val Gly Pro
 130 135 140
 Thr Asn Pro Glu Ser Pro Pro Leu Gln Ser Pro Ala Pro Pro Ala
 145 150 155 160
 Ser Asp Pro Thr Asn Ser Pro Pro Ala Ser Pro Leu Asp Pro Thr Asn
 165 170 175
 Pro Pro Pro Ile Gln Pro Ser Gly Pro Ala Thr Ser Pro Pro Ala Asn
 180 185 190
 Pro Asn Ala Pro Pro Ser Pro Phe Pro Thr Val Pro Pro Lys Thr Pro
 195 200 205
 Ser Ser Gly Pro Val Val Ser Pro Ser Leu Thr Ser Pro Ser Lys Gly
 210 215 220
 Thr Pro Thr Pro Asn Gln Gly Asn Gly Asp Gly Gly Gly Gly Gly Gly
 225 230 235 240
 Gly Tyr Gln Gly Lys Thr Met Val Gly Met Ala Val Ala Gly Phe Ala
 245 250 255
 Ile Met Ala Leu Ile Gly Val Val Phe Leu Val Arg Arg Lys Lys
 260 265 270
 Arg Asn Ile Asp Ser Tyr Asn His Ser Gln Tyr Leu Pro His Pro Asn
 275 280 285
 Phe Ser Val Lys Ser Asp Gly Phe Leu Tyr Gly Gln Asp Pro Gly Lys

ALL INFORMATION CONTAINED HEREIN IS UNCLASSIFIED

	290					295					300					
Gly 305	Tyr	Ser	Ser	Gly	Pro 310	Asn	Gly	Ser	Met	Tyr 315	Asn	Asn	Ser	Gln	Gln 320	
Gln	Gln	Ser	Ser	Met 325	Gly	Asn	Ser	Tyr	Gly 330	Thr	Ala	Gly	Gly	Gly 335	Tyr	
Pro	His	His	Gln 340	Met	Gln	Ser	Ser	Gly 345	Thr	Pro	Asp	Ser	Ala 350	Ile	Leu	
Gly	Ser	Gly 355	Gln	Thr	His	Phe	Ser 360	Tyr	Glu	Glu	Leu	Ala 365	Glu	Ile	Thr	
Gln	Gly 370	Phe	Ala	Arg	Lys	Asn 375	Ile	Leu	Gly	Glu	Gly 380	Gly	Phe	Gly	Cys	
Val 385	Tyr	Lys	Gly	Thr	Leu 390	Gln	Asp	Gly	Lys	Val 395	Val	Ala	Val	Lys	Gln 400	
Leu	Lys	Ala	Gly	Ser 405	Gly	Gln	Gly	Asp	Arg 410	Glu	Phe	Lys	Ala 415	Glu	Val	
Glu	Ile	Ile	Ser 420	Arg	Val	His	His	Arg 425	His	Leu	Val	Ser	Leu 430	Val	Gly	
Tyr	Cys	Ile 435	Ser	Asp	Gln	His	Arg 440	Leu	Leu	Ile	Tyr	Glu 445	Tyr	Val	Ser	
Asn	Gln 450	Thr	Leu	Glu	His	His 455	Leu	His	Glu	Trp	Ser	Lys	Arg	Val	Arg	
Ile 465	Ala	Ile	Gly	Ser	Ala 470	Lys	Gly	Leu	Ala	Tyr 475	Leu	His	Glu	Asp	Cys 480	
His	Pro	Lys	Ile 485	Ile	His	Arg	Asp	Ile	Lys 490	Ser	Ala	Asn	Ile	Leu 495	Leu	
Asp	Asp	Glu	Tyr 500	Glu	Ala	Gln	Ala	Ile 505	Met	Lys	Ser	Ser	Phe 510	Ser	Leu	
Asn	Leu	Ser 515	Tyr	Asp	Cys	Lys	Val 520	Leu	Val	Ala	Asp	Phe 525	Gly	Leu	Ala	
Arg	Leu 530	Asn	Asp	Thr	Thr	Gln 535	Thr	His	Val	Ser	Thr 540	Arg	Val	Met	Gly	
Thr 545	Phe	Gly	Tyr	Leu	Ala 550	Pro	Glu	Tyr	Ala	Ser 555	Ser	Gly	Lys	Leu	Thr 560	
Asp	Arg	Ser	Asp 565	Val	Phe	Ser	Phe	Gly	Val 570	Val	Leu	Leu	Glu	Leu 575	Val	
Thr	Gly	Arg	Lys 580	Pro	Val	Asp	Gln	Thr 585	Gln	Pro	Leu	Gly	Glu 590	Glu	Ser	
Leu	Val	Glu 595	Trp	Ala	Arg	Pro	Leu 600	Leu	Leu	Lys	Ala	Ile 605	Glu	Thr	Gly	
Asp	Leu 610	Ser	Glu	Leu	Ile	Asp 615	Thr	Arg	Leu	Glu	Lys 620	Arg	Tyr	Val	Glu	
His 625	Glu	Val	Phe	Arg	Met 630	Ile	Glu	Thr	Ala	Ala 635	Ala	Cys	Val	Arg	His 640	
Ser	Gly	Pro	Lys 645	Arg	Pro	Arg	Met	Val	Gln 650	Val	Val	Arg	Ala	Leu 655	Asp	
Cys	Asp	Gly 660	Asp	Ser	Gly	Asp	Ile	Ser 665	Asn	Gly	Ile	Lys	Ile 670	Gly	Gln	
Ser	Thr	Thr 675	Tyr	Asp	Ser	Gly	Gln 680	Tyr	Asn	Glu	Asp	Ile 685	Met	Lys	Phe	
Arg	Lys 690	Met	Ala	Phe	Gly	Gly 695	Asp	Asn	Ser	Val	Glu 700	Ser	Gly	Leu	Tyr	
Ser 705	Gly	Asn	Tyr	Ser	Ala 710	Lys	Ser	Ser	Ser	Asp 715	Phe	Ser	Gly	Asn	Glu 720	
Ser	Glu	Thr	Arg 725	Pro	Phe	Asn	Asn	Arg 730	Arg	Phe						

THE UNIVERSITY OF CHICAGO

<400>	9					
gaaaatttttg	atctccgatg	gtttcttctc	ctgaatctgc	tcctccaaca	aactccacct	60
cttctccatc	tccaccgtct	aataccaatt	caaccacctc	ttctccgccg	gtcccgctct	120
ctcctttctc	tacacctcct	caaggagact	catcatcatc	gccacctcct	gattccacat	180
ctccaccagc	tccacaagct	cctaaccctc	ctaattctct	taataactct	ccttcccctc	240
cgtcacaggg	cgggtggagga	gaaagaggaa	atggaggaaa	caatggtggc	aatgatactc	300
caccgtcacg	cggctctcct	ccttctctct	cttctaggag	taatggagat	aatggtggta	360
gcagatcatc	gccaccagga	gacactggag	gctctcgctc	agacaacctc	ccttctagcg	420
gaggaagcag	tggaggagggt	ggaggtggaa	gaagtaatac	gaatacagcg	atcatagttg	480
gtgtattagt	cggagcttga	cttttgatga	tcgttcttat	tattgtgtgt	cttagacgca	540
aaaagaagag	aaaagactcc	ttctaccctg	aacccatgaa	aggaaaccaa	tatcaatact	600
atggaaacaa	caacaacaac	aatgtctcac	agaattatcc	gaattggcac	ctaaattcac	660
aaggccaaaa	ccaacaatct	actggtggtt	ggggaggcgg	tggaccatca	ccgcctcctc	720
ctccgcggat	gcctacaagc	ggagaagatt	cttccatgta	ctcaggccca	tcacgcccag	780
ttttacctcc	tccttcgcct	gctctagccc	tcggattcaa	caagagcact	tttacttacc	840
aagagcttgc	ggctgcaaca	ggagggttta	cggatgctaa	ccttttggga	cagggaggat	900
ttgggtatgt	ccataaagga	gtcttgctta	gcgggaaaga	agtagcagtt	aagagtttaa	960
aagcgggtag	cggacaagga	gagagggagt	ttcaagctga	ggtcgatatc	attagccgtg	1020
tgcattcatcg	gtatcttgtt	tctttggttg	gatattgcac	agctgatgga	cagaggatgt	1080
tggtttatga	gtttgttctt	aacaaaaact	tggaatatca	tcttcatggg	aaaaatcttc	1140
cggtaatatga	gttctccact	aggttgcgta	tcgccttagg	tgctgcgaaa	ggactcgctt	1200
accttcacga	agactgccat	cctcggatca	ttcacccgca	catcaagtct	gcaaatatct	1260
tcttggaactt	caactttgat	gctatggtgg	ctgatttttg	attagctaag	ttaacatctg	1320
ataacaacac	tcattgtatct	actcgtgtga	tgggaacttt	cggatatcta	gctccagaat	1380
atgcttcaag	cggtaaatta	accgagaaat	ccgatgtttt	ctcttacgga	gttatgttat	1440
tggaaacttat	aactggaaaa	cgaccgggtt	ataatagcat	caccatggac	gacaccttag	1500
tagattgggc	tcggcctctt	atggctcgcg	cgctagaaga	tggaaacttt	aatgagctcg	1560
cagatgcgag	gcttgaaggc	aactacaacc	cgcaagaaat	ggctcgaatg	gtgacttgtg	1620
ccgctgctag	cattcgctcat	tcggggcgta	aacgtccaaa	gatgagccag	atagtaagag	1680
cgttagaagg	agaagtgtcc	ttagatgctt	taaacgaagg	tgtgaagcca	ggacacagta	1740
acgtttacgg	gtcattggga	gcaagctcgg	attatagtca	gacatcttac	aatgcagaca	1800
tgaagaattt	cagacagata	gctttgtcga	gccaaagaatt	cccagtcagt	gactgtgaag	1860
gaacatctag	taatgatctt	agagatatgg	gaactaagag	ccctactcct	ccaaaatgag	1920
atcgaatcaa	tqattctgt					1939

```
<210> 10
<211> 1902
<212> DNA
<213> Arabidopsis thaliana
```

```
<220>
<221> CDS
<222> (1) .. (1902)
```

<400> 10																
atg	gct	tct	tct	cct	gaa	tct	gct	cct	cca	aca	aac	tcc	acc	tct	tct	48
Met	Ala	Ser	Ser	Pro	Glu	Ser	Ala	Pro	Pro	Thr	Asn	Ser	Thr	Ser	Ser	
1				5					10					15		
cca	tct	cca	ccg	tct	aat	acc	aat	tca	acc	acc	tct	tct	ccg	ccg	gct	96
Pro	Ser	Pro	Pro	Ser	Asn	Thr	Asn	Ser	Thr	Thr	Ser	Ser	Pro	Pro	Ala	
			20					25					30			

ccg	tct	cct	cct	tct	cct	aca	cct	cct	caa	gga	gac	tca	tca	tca	tcg	144
Pro	Ser	Pro	Pro	Ser	Pro	Thr	Pro	Pro	Gln	Gly	Asp	Ser	Ser	Ser	Ser	
		35					40				45					
cca	cct	cct	gat	tcc	aca	tct	cca	cca	gct	cca	caa	gct	cct	aac	cct	192
Pro	Pro	Pro	Asp	Ser	Thr	Ser	Pro	Pro	Ala	Pro	Gln	Ala	Pro	Asn	Pro	
	50					55					60					
cct	aat	tcc	tct	aat	aac	tct	cct	tcc	cct	ccg	tca	cag	ggc	ggt	gga	240
Pro	Asn	Ser	Ser	Asn	Asn	Ser	Pro	Ser	Pro	Pro	Ser	Gln	Gly	Gly	Gly	
65					70					75					80	
gga	gaa	aga	gga	aat	gga	gga	aac	aat	ggt	ggc	aat	gat	act	cca	ccg	288
Gly	Glu	Arg	Gly	Asn	Gly	Gly	Asn	Asn	Gly	Gly	Asn	Asp	Thr	Pro	Pro	
				85					90					95		
tca	cgc	ggc	tct	cct	cct	tct	cct	cct	tct	agg	agt	aat	gga	gat	aat	336
Ser	Arg	Gly	Ser	Pro	Pro	Ser	Pro	Pro	Ser	Arg	Ser	Asn	Gly	Asp	Asn	
			100					105					110			
ggt	ggt	agc	aga	tca	tcg	cca	cca	gga	gac	act	gga	ggc	tct	cgc	tca	384
Gly	Gly	Ser	Arg	Ser	Ser	Pro	Pro	Gly	Asp	Thr	Gly	Gly	Ser	Arg	Ser	
		115					120					125				
gac	aac	cct	cct	tct	agc	gga	gga	agc	agt	gga	gga	ggt	gga	ggt	gga	432
Asp	Asn	Pro	Pro	Ser	Ser	Gly	Gly	Ser	Ser	Gly	Gly	Gly	Gly	Gly	Gly	
	130					135					140					
aga	agt	aat	acg	aat	aca	gcg	atc	ata	gtt	ggt	gta	tta	gtc	gga	gct	480
Arg	Ser	Asn	Thr	Asn	Thr	Ala	Ile	Ile	Val	Gly	Val	Leu	Val	Gly	Ala	
145					150					155					160	
gga	ctt	ttg	atg	atc	gtt	ctt	att	att	gtg	tgt	ctt	aga	cgc	aaa	aag	528
Gly	Leu	Leu	Met	Ile	Val	Leu	Ile	Ile	Val	Cys	Leu	Arg	Arg	Lys	Lys	
				165					170					175		
aag	aga	aaa	gac	tcc	ttc	tac	cct	gaa	ccc	atg	aaa	gga	aac	caa	tat	576
Lys	Arg	Lys	Asp	Ser	Phe	Tyr	Pro	Glu	Pro	Met	Lys	Gly	Asn	Gln	Tyr	
			180					185					190			
caa	tac	tat	gga	aac	aac	aac	aac	aac	aat	gct	tca	cag	aat	tat	ccg	624
Gln	Tyr	Tyr	Gly	Asn	Asn	Asn	Asn	Asn	Asn	Ala	Ser	Gln	Asn	Tyr	Pro	
		195					200					205				
aat	tgg	cac	cta	aat	tca	caa	ggc	caa	aac	caa	caa	tct	act	ggt	ggt	672
Asn	Trp	His	Leu	Asn	Ser	Gln	Gly	Gln	Asn	Gln	Gln	Ser	Thr	Gly	Gly	
	210					215					220					
tgg	gga	ggc	ggt	gga	cca	tca	ccg	cct	cct	cct	ccg	cgg	atg	cct	aca	720
Trp	Gly	Gly	Gly	Gly	Pro	Ser	Pro	Pro	Pro	Pro	Pro	Arg	Met	Pro	Thr	
225					230					235					240	
agc	gga	gaa	gat	tct	tcc	atg	tac	tca	ggc	cca	tca	cgc	cca	gtt	tta	768
Ser	Gly	Glu	Asp	Ser	Ser	Met	Tyr	Ser	Gly	Pro	Ser	Arg	Pro	Val	Leu	
				245					250					255		

cct Pro	cct Pro	cct Pro	tcg Ser 260	cct Pro	gct Ala	cta Leu	gcc Ala	ctc Leu 265	gga Gly	ttc Phe	aac Asn	aag Lys	agc Ser 270	act Thr	ttt Phe	816
act Thr	tac Tyr	caa Gln 275	gag Glu	ctt Leu	gcg Ala	gct Ala	gca Ala 280	aca Thr	gga Gly	ggg Gly	ttt Phe	acg Thr 285	gat Asp	gct Ala	aac Asn	864
ctt Leu 290	ttg Leu	gga Gly	cag Gln	gga Gly	gga Gly	ttt Phe 295	ggg Gly	tat Tyr	gtc Val	cat His	aaa Lys 300	gga Gly	gtc Val	ttg Leu	cct Pro	912
agc Ser 305	ggg Gly	aaa Lys	gaa Glu	gta Val	gca Ala 310	gtt Val	aag Lys	agt Ser	tta Leu	aaa Lys 315	gcg Ala	ggt Gly	agc Ser	gga Gly	caa Gln 320	960
gga Gly	gag Glu	agg Arg	gag Glu	ttt Phe 325	caa Gln	gct Ala	gag Glu	gtc Val	gat Asp 330	atc Ile	att Ile	agc Ser	cgt Arg	gtg Val 335	cat His	1008
cat His	cgg Arg	tat Tyr	ctt Leu 340	gtt Val	tct Ser	ttg Leu	gtt Val	gga Gly 345	tat Tyr	tgc Cys	ata Ile	gct Ala	gat Asp 350	gga Gly	cag Gln	1056
agg Arg	atg Met	ttg Leu 355	gtt Val	tat Tyr	gag Glu	ttt Phe	gtt Val 360	cct Pro	aac Asn	aaa Lys	act Thr	ttg Leu 365	gaa Glu	tat Tyr	cat His	1104
ctt Leu 370	cat His	ggg Gly	aaa Lys	aat Asn	ctt Leu	ccg Pro 375	gta Val	atg Met	gag Glu	ttc Phe	tcc Ser 380	act Thr	agg Arg	ttg Leu	cgt Arg	1152
atc Ile 385	gcc Ala	tta Leu	ggt Gly	gct Ala	gcg Ala 390	aaa Lys	gga Gly	ctc Leu	gct Ala	tac Tyr 395	ctt Leu	cac His	gaa Glu	gac Asp	tgc Cys 400	1200
cat His	cct Pro	cgg Arg	atc Ile	att Ile 405	cac His	cgc Arg	gac Asp	atc Ile	aag Lys 410	tct Ser	gca Ala	aat Asn	att Ile	ctc Leu 415	ttg Leu	1248
gac Asp	ttc Phe	aac Asn	ttt Phe 420	gat Asp	gct Ala	atg Met	gtg Val	gct Ala 425	gat Asp	ttt Phe	gga Gly	tta Leu	gct Ala 430	aag Lys	tta Leu	1296
aca Thr	tct Ser	gat Asp 435	aac Asn	aac Asn	act Thr	cat His	gta Val 440	tct Ser	act Thr	cgt Arg	gtg Val	atg Met 445	gga Gly	act Thr	ttc Phe	1344
gga Gly	tat Tyr 450	cta Leu	gct Ala	cca Pro	gaa Glu	tat Tyr 455	gct Ala	tca Ser	agc Ser	ggg Gly	aaa Lys 460	tta Leu	acc Thr	gag Glu	aaa Lys	1392
tcc Ser 465	gat Asp	gtt Val	ttc Phe	tct Ser	tac Tyr 470	gga Gly	gtt Val	atg Met	tta Leu	ttg Leu 475	gaa Glu	ctt Leu	ata Ile	act Thr	gga Gly 480	1440
aaa Lys	cga Arg	ccg Pro	gtt Val	gat Asp	aat Asn	agc Ser	atc Ile	acc Thr	atg Met	gac Asp	gac Asp	acc Thr	tta Leu	gta Val	gat Asp	1488

Ser Arg Gly Ser Pro Pro Ser Pro Pro Ser Arg Ser Asn Gly Asp Asn

				565						570					575				
Tyr	Gly	Ser	Leu	Gly	Ala	Ser	Ser	Asp	Tyr	Ser	Gln	Thr	Ser	Tyr	Asn				
			580					585					590						
Ala	Asp	Met	Lys	Lys	Phe	Arg	Gln	Ile	Ala	Leu	Ser	Ser	Gln	Glu	Phe				
		595					600					605							
Pro	Val	Ser	Asp	Cys	Glu	Gly	Thr	Ser	Ser	Asn	Asp	Ser	Arg	Asp	Met				
	610					615					620								
Gly	Thr	Lys	Ser	Pro	Thr	Pro	Pro	Lys											
625					630														

<210> 12
 <211> 2104
 <212> DNA
 <213> Arabidopsis thaliana

<400> 12

tccaccgttt	gagaaaccct	aataacaaca	ttcaaaatgg	cggactcacc	ggtggattca	60
tctcctgccc	ctgaaaccct	aaatgggaca	ccaccgtcaa	acggaacatc	gccgtctaata	120
gagtcacgc	cgccaacacc	accttcttca	ccaccaccat	catcaatata	tgctcctccg	180
ccagatatct	ccgcttcttt	ttcaccgcgc	cctgcaccac	caacgcaaga	aacgtcacct	240
cctacatctc	cgctcctcatc	gccgcctggt	gtagctaata	cgtcaccgca	gactccagag	300
aatccttctc	cacctgcacc	tgaaggctca	actcctgtaa	cgccacctgc	accaccacaa	360
acaccgtcga	accaatcacc	ggaaagacca	actcctcctt	ctcctgggtgc	caatgatgac	420
cgaaacagaa	ccaatggcgg	aaacaacaac	agagacggct	ccacaccatc	accaccgtcg	480
tcagggaaca	gaacttccgg	tgacgggtggc	tcaccttcac	cacctcgggtc	gataagccct	540
cctcagaata	gtggagattc	agactcatca	tcggggcctt	tgcttctact	tgacgtgtgt	600
atttgcattc	gttgcaacag	gaagaagaag	aagaaatctc	ctcagggtcaa	ccacatgcac	660
tactacaata	acaatcctta	tgaggagagca	ccctcaggta	atggtgggtta	ttacaaggga	720
acacctcaag	atcatgtggt	gaatatggct	gggtcaaggag	gtgggaattg	gggtccacag	780
caacctgtgt	ctggctctca	cagtgatgct	tccaacttaa	cgggtcgaac	tgctataccg	840
tcacctcaag	ctgcaactct	tggtcacaac	caaagcactt	tcacatacga	tgaactgtcc	900
attgcaacag	aagggttctgc	tcagtcaaata	ttgctaggac	aaggaggatt	tggttatgtt	960
cataaaggag	ttctgcctag	tggaacaagaa	gttgacagtga	agagtcttaa	acttgggaagt	1020
ggacaagggg	aacgcgagtt	tcaagcagag	gttgatatca	ttagccgtgt	ccatcatcgt	1080
catctcgttt	ctcttggttg	atattgcata	tctggtggtc	aaagactttt	ggttttatgag	1140
tttataccta	acaacactct	tgaatttcat	cttcatggaa	agggtcgtcc	ggtttttggat	1200
tggcctacaa	gagtgaagat	tgcatatggga	tcagctagag	gccttgcata	tttgcataaa	1260
gactgtaaga	aaatctttat	ctcacatatt	tgcatcagtc	accctcgcac	tatccacaga	1320
gatatcaaa	ctgcaaaccat	tcttcttgat	ttcagttttg	agaccaagggt	ggcagatttt	1380
ggattggcta	agctatctca	agacaactat	actcatgtct	ccactcgcgt	catgggaact	1440
tttggtatact	tagctccaga	gtatgcata	agcggaaagt	tatccgacaa	atctgatgtt	1500
ttctcatttg	gagtaatgct	tcttgagctc	ataaccggaa	gacctcctct	ggatctaact	1560
ggagaaatgg	aagatagctt	ggtagattgg	gcaaggcctt	tgtgtttgaa	agcagctcaa	1620
gatggagatt	acaaccaatt	ggctgatcca	cgtctagagc	taaaactacag	tcatcaagag	1680
atggttcaaa	tggtctcttg	tgacagctga	gcaatcagac	attcagcaag	aagacggcct	1740
aagatgagcc	aggttcaaaa	actcatacca	cttggttggtt	ctattattgt	acgagcacta	1800
gaaggagata	tgtcaatgga	tgatctaagt	gagggaaaca	gaccaggaca	aagcacgtac	1860
ttgagccccg	ggagcgtgag	ctcagagtat	gacgcaagct	cgtacacggc	agacatgaaa	1920
aagttcaaga	aactggcggt	agagaataaa	gaatatcaaa	gcagtgaata	tggtggaaca	1980
agtgagtatg	gcttaaacc	ttctgcttca	agtagtgaag	aaatgaatag	aggctcaatg	2040
aaacgcaatc	ctcagctttg	aaagaagaga	caacacttgt	cataatattt	cagttttctt	2100
ctct						2104

<210> 13
 <211> 2025

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(2025)

<400> 13

atg	gcg	gac	tca	ccg	gtg	gat	tca	tct	cct	gcc	cct	gaa	acc	tca	aat	48
Met	Ala	Asp	Ser	Pro	Val	Asp	Ser	Ser	Pro	Ala	Pro	Glu	Thr	Ser	Asn	
1				5					10					15		
ggg	aca	cca	ccg	tca	aac	gga	aca	tcg	ccg	tct	aat	gag	tca	tcg	ccg	96
Gly	Thr	Pro	Pro	Ser	Asn	Gly	Thr	Ser	Pro	Ser	Asn	Glu	Ser	Ser	Pro	
			20					25					30			
cca	aca	cca	cct	tct	tca	cca	cca	cca	tca	tca	ata	tct	gct	cct	ccg	144
Pro	Thr	Pro	Pro	Ser	Ser	Pro	Pro	Pro	Ser	Ser	Ile	Ser	Ala	Pro	Pro	
			35					40				45				
cca	gat	atc	tcc	gct	tct	ttt	tca	ccg	ccg	cct	gca	cca	cca	acg	caa	192
Pro	Asp	Ile	Ser	Ala	Ser	Phe	Ser	Pro	Pro	Pro	Ala	Pro	Pro	Thr	Gln	
	50					55					60					
gaa	acg	tca	cct	cct	aca	tct	ccg	tcc	tca	tcg	ccg	cct	gtt	gta	gct	240
Glu	Thr	Ser	Pro	Pro	Thr	Ser	Pro	Ser	Ser	Ser	Pro	Pro	Val	Val	Ala	
65					70					75					80	
aat	ccg	tca	ccg	cag	act	cca	gag	aat	cct	tct	cca	cct	gca	cct	gaa	288
Asn	Pro	Ser	Pro	Gln	Thr	Pro	Glu	Asn	Pro	Ser	Pro	Pro	Ala	Pro	Glu	
				85					90					95		
ggc	tca	act	cct	gta	acg	cca	cct	gca	cca	cca	caa	aca	ccg	tcg	aac	336
Gly	Ser	Thr	Pro	Val	Thr	Pro	Pro	Ala	Pro	Pro	Gln	Thr	Pro	Ser	Asn	
			100					105					110			
caa	tca	ccg	gaa	aga	cca	act	cct	cct	tct	cct	ggt	gcc	aat	gat	gac	384
Gln	Ser	Pro	Glu	Arg	Pro	Thr	Pro	Pro	Ser	Pro	Gly	Ala	Asn	Asp	Asp	
		115					120					125				
cga	aac	aga	acc	aat	ggc	gga	aac	aac	aac	aga	gac	ggc	tcc	aca	cca	432
Arg	Asn	Arg	Thr	Asn	Gly	Gly	Asn	Asn	Asn	Arg	Asp	Gly	Ser	Thr	Pro	
	130					135					140					
tca	cca	ccg	tcg	tca	ggg	aac	aga	act	tcc	ggt	gac	ggt	ggc	tca	cct	480
Ser	Pro	Pro	Ser	Ser	Gly	Asn	Arg	Thr	Ser	Gly	Asp	Gly	Gly	Ser	Pro	
145					150					155					160	
tca	cca	cct	cgg	tcg	ata	agc	cct	cct	cag	aat	agt	gga	gat	tca	gac	528
Ser	Pro	Pro	Arg	Ser	Ile	Ser	Pro	Pro	Gln	Asn	Ser	Gly	Asp	Ser	Asp	
				165					170					175		
tca	tca	tcg	ggg	ctt	ttg	ctt	cta	ctt	gca	gtg	tgt	att	tgc	atc	tgt	576
Ser	Ser	Ser	Gly	Leu	Leu	Leu	Leu	Leu	Ala	Val	Cys	Ile	Cys	Ile	Cys	
			180					185					190			
tgc	aac	agg	aag	aag	aag	aag	aaa	tct	cct	cag	gtc	aac	cac	atg	cac	624
Cys	Asn	Arg	Lys	Lys	Lys	Lys	Lys	Ser	Pro	Gln	Val	Asn	His	Met	His	

tac Tyr	tac Tyr 210	aat Asn	aac Asn	aat Asn	cct Pro	tat Tyr 215	gga Gly	gga Gly	gca Ala	ccc Pro	tca Ser 220	ggt Gly	aat Asn	ggt Gly	ggt Gly	672
tat Tyr 225	tac Tyr	aag Lys	gga Gly	aca Thr	cct Pro 230	caa Gln	gat Asp	cat His	gtg Val	gtg Val 235	aat Asn	atg Met	gct Ala	ggt Gly	caa Gln 240	720
gga Gly	ggt Gly	ggg Gly	aat Asn	tgg Trp 245	ggt Gly	cca Pro	cag Gln	caa Gln	cct Pro 250	gtg Val	tct Ser	ggt Gly	cct Pro	cac His 255	agt Ser	768
gat Asp	gct Ala	tcc Ser	aac Asn 260	tta Leu	acc Thr	ggt Gly	cga Arg	act Thr 265	gct Ala	ata Ile	ccg Pro	tca Ser	cct Pro 270	caa Gln	gct Ala	816
gca Ala	act Thr	ctt Leu 275	ggt Gly	cac His	aac Asn	caa Gln	agc Ser 280	act Thr	ttc Phe	aca Thr	tac Tyr	gat Asp 285	gaa Glu	ctg Leu	tcc Ser	864
att Ile	gca Ala 290	aca Thr	gaa Glu	ggt Gly	ttc Phe	gct Ala 295	cag Gln	tca Ser	aat Asn	ttg Leu 300	cta Leu	gga Gly	caa Gln	gga Gly	gga Gly	912
ttt Phe 305	ggg Gly	tat Tyr	gtt Val	cat His	aaa Lys 310	gga Gly	gtt Val	ctg Leu	cct Pro	agt Ser 315	ggc Gly	aaa Lys	gaa Glu	gtt Val	gca Ala 320	960
gtg Val	aag Lys	agt Ser	ctt Leu	aaa Lys 325	ctt Leu	gga Gly	agt Ser	gga Gly	caa Gln 330	ggg Gly	gaa Glu	cgc Arg	gag Glu	ttt Phe 335	caa Gln	1008
gca Ala	gag Glu	gtt Val	gat Asp 340	atc Ile	att Ile	agc Ser	cgt Arg	gtc Val 345	cat His	cat His	cgt Arg	cat His	ctc Leu 350	gtt Val	tct Ser	1056
ctt Leu	gtt Val	gga Gly 355	tat Tyr	tgc Cys	atc Ile	tct Ser	ggt Gly 360	ggt Gly	caa Gln	aga Arg	ctt Leu	ttg Leu 365	gtt Val	tat Tyr	gag Glu	1104
ttt Phe	ata Ile 370	cct Pro	aac Asn	aac Asn	act Thr	ctt Leu 375	gaa Glu	ttt Phe	cat His	ctt Leu	cat His 380	gga Gly	aag Lys	ggt Gly	cgt Arg	1152
ccg Pro 385	gtt Val	ttg Leu	gat Asp	tgg Trp	cct Pro 390	aca Thr	aga Arg	gtg Val	aag Lys	att Ile 395	gca Ala	ttg Leu	gga Gly	tca Ser	gct Ala 400	1200
aga Arg	ggc Gly	ctt Leu	gca Ala	tat Tyr 405	ttg Leu	cat His	gaa Glu	gac Asp	tgt Cys 410	aag Lys	aaa Lys	atc Ile	ttt Phe	atc Ile 415	tca Ser	1248
cat His	att Ile	tgc Cys	atc Ile 420	agt Ser	cac His	cct Pro	cgc Arg	att Ile 425	atc Ile	cac His	aga Arg	gat Asp	atc Ile 430	aaa Lys	gct Ala	1296

— 222 —

660

665

670

cag ctt tga
Gln Leu

2025

<210> 14
<211> 674
<212> PRT
<213> Arabidopsis thaliana

<400> 14

Met	Ala	Asp	Ser	Pro	Val	Asp	Ser	Ser	Pro	Ala	Pro	Glu	Thr	Ser	Asn
1				5					10					15	
Gly	Thr	Pro	Pro	Ser	Asn	Gly	Thr	Ser	Pro	Ser	Asn	Glu	Ser	Ser	Pro
		20						25					30		
Pro	Thr	Pro	Pro	Ser	Ser	Pro	Pro	Pro	Ser	Ser	Ile	Ser	Ala	Pro	Pro
		35					40					45			
Pro	Asp	Ile	Ser	Ala	Ser	Phe	Ser	Pro	Pro	Pro	Ala	Pro	Pro	Thr	Gln
	50					55					60				
Glu	Thr	Ser	Pro	Pro	Thr	Ser	Pro	Ser	Ser	Ser	Pro	Pro	Val	Val	Ala
65					70					75					80
Asn	Pro	Ser	Pro	Gln	Thr	Pro	Glu	Asn	Pro	Ser	Pro	Pro	Ala	Pro	Glu
				85					90					95	
Gly	Ser	Thr	Pro	Val	Thr	Pro	Pro	Ala	Pro	Pro	Gln	Thr	Pro	Ser	Asn
			100					105					110		
Gln	Ser	Pro	Glu	Arg	Pro	Thr	Pro	Pro	Ser	Pro	Gly	Ala	Asn	Asp	Asp
	115						120					125			
Arg	Asn	Arg	Thr	Asn	Gly	Gly	Asn	Asn	Asn	Arg	Asp	Gly	Ser	Thr	Pro
	130				135						140				
Ser	Pro	Pro	Ser	Ser	Gly	Asn	Arg	Thr	Ser	Gly	Asp	Gly	Gly	Ser	Pro
145					150					155					160
Ser	Pro	Pro	Arg	Ser	Ile	Ser	Pro	Pro	Gln	Asn	Ser	Gly	Asp	Ser	Asp
			165						170					175	
Ser	Ser	Ser	Gly	Leu	Leu	Leu	Leu	Leu	Ala	Val	Cys	Ile	Cys	Ile	Cys
			180					185					190		
Cys	Asn	Arg	Lys	Lys	Lys	Lys	Lys	Ser	Pro	Gln	Val	Asn	His	Met	His
	195						200					205			
Tyr	Tyr	Asn	Asn	Asn	Pro	Tyr	Gly	Gly	Ala	Pro	Ser	Gly	Asn	Gly	Gly
	210					215					220				
Tyr	Tyr	Lys	Gly	Thr	Pro	Gln	Asp	His	Val	Val	Asn	Met	Ala	Gly	Gln
225					230					235					240
Gly	Gly	Gly	Asn	Trp	Gly	Pro	Gln	Gln	Pro	Val	Ser	Gly	Pro	His	Ser
			245						250					255	
Asp	Ala	Ser	Asn	Leu	Thr	Gly	Arg	Thr	Ala	Ile	Pro	Ser	Pro	Gln	Ala
			260					265					270		
Ala	Thr	Leu	Gly	His	Asn	Gln	Ser	Thr	Phe	Thr	Tyr	Asp	Glu	Leu	Ser
		275					280					285			
Ile	Ala	Thr	Glu	Gly	Phe	Ala	Gln	Ser	Asn	Leu	Leu	Gly	Gln	Gly	Gly
	290					295					300				
Phe	Gly	Tyr	Val	His	Lys	Gly	Val	Leu	Pro	Ser	Gly	Lys	Glu	Val	Ala
305					310					315					320
Val	Lys	Ser	Leu	Lys	Leu	Gly	Ser	Gly	Gln	Gly	Glu	Arg	Glu	Phe	Gln
			325						330					335	
Ala	Glu	Val	Asp	Ile	Ile	Ser	Arg	Val	His	His	Arg	His	Leu	Val	Ser
			340					345					350		
Leu	Val	Gly	Tyr	Cys	Ile	Ser	Gly	Gly	Gln	Arg	Leu	Leu	Val	Tyr	Glu


```

tttctctctc ccatcaattc ttctccacca aatccttctc cgaacactcc gtcactccca 600
gaaacttctc ctccacctaa accaccgctc tcaacgacgc catttccctc ctcatccact 660
cccccgcta agaagtcctc tgcagcagta actcttcctt tctttggggc agcggggccaa 720
ttaccggatg ggaccgtagc acctcctatt gggcctgtta tagagcccaa gacgagtcca 780
gccgaatcaa tatctccggg aacgccacag ccactgggtc cgaagagtct acctgtaacg 840
acgtcgtatc accgatcatc cgccggattc ttatttggcg gtgtaatcgt tggagctctt 900
ctactaattc tgtttaggtct tctctttgtc ttctacagag ctaccagaaa tagaataaac 960
aacagcagct ctgctcatca tcaatccaaa actccctcaa aagttcaaca tcatcggggc 1020
ggtaatgctg gtacgaacca ggcacatggt atcacaatgc caccaccaat ccatgctaaa 1080
tatatatcta gtggagggtg tgatacgaag gagaacaatt ctgttgcgaa aaacatttca 1140
atgccatctg gaatgttctc ctacgaagaa ctttcaaaag caactgggtg attttcagag 1200
gagaaccttt tgggagaagg cggtttcgga tatgttcaca aaggagtgtt gaaaaacggg 1260
acagaagttg cggatgaagca gctgaagatt gggagctatc aaggggaaag agaattccaa 1320
gctgaggttg acacaatcag tagggttcat cataagcacc tcgttttcatt ggttgggttat 1380
tgcgttaatg gagataaaaag actcttgggt tacgagtttg ttcctaaaga taccttggag 1440
ttccacttgc atgagaacag aggaagcgtg ttggaatggg aaatgaggct caggattgct 1500
gtaggagcag caaaaggatt agcttatctt catgaggatt gcagtccaac tataattcac 1560
cgtgatatca aagcagctaa tatccttcta gattccaaat ttgaggcaaa ggtctctgac 1620
tttgactag ccaagttttt ctccagacac aattcatcat tcatcatat ctctactcga 1680
gtggtaggaa ctttcggata catggctcca gaatacgcgt ccagtggtaa agtaactgat 1740
aaatcagatg tatattcctt tggggctcgt cttctagaac tcatcactgg acgtccatca 1800
attttcgcca aagattcttc cacaaaccag agtttagtag actgggagag gccattgctt 1860
acgaaagcaa tctctggaga aagttttgac tttctttag actcaagggt ggagaagaat 1920
tacgatacaa ctccagatggc aaacatggct gcttgtgctg ctgcttgcac acgccaatca 1980
gcttggcttc ggcctagaat gagccaggta gtacgtgctc ttgaaggcga ggtggccctg 2040
agaaaggctg aagagactgg gaatagcgtg acctatagct cttctgaaaa cccgaatgac 2100
atcacaccac ggtatggaac aaataagagg agattcgaca caggttcaag cgatgggttac 2160
acttcagaat atggagttaa cccttctcag tcgagcagtg aacatcaaca ggtgaatact 2220
tagttcacag gttcaatagg gcaagtttca ccacaattat t 2261

```

<210> 16

<211> 2196

<212> DNA

<213> Arabidopsis thaliana

<220>

<221> CDS

<222> (1)..(2196)

<400> 16

```

atg tcg ctc tct cct tct tcc tcg ccg gct ccg gca act tcc ccg cct 48
Met Ser Leu Ser Pro Ser Ser Ser Pro Ala Pro Ala Thr Ser Pro Pro
1 5 10 15

```

```

gcc atg tca tta cct ccg gca gat tcc gta cct gac acg tca tca cct 96
Ala Met Ser Leu Pro Pro Ala Asp Ser Val Pro Asp Thr Ser Ser Pro
20 25 30

```

```

cca gct cct cct ttg tct cct ctt ccc cca cca ttg agc tct cct ccg 144
Pro Ala Pro Pro Leu Ser Pro Leu Pro Pro Pro Leu Ser Ser Pro Pro
35 40 45

```

```

ccg ttg cct tca cca ccg cct ctc tcc gct ccc acc gct tcc cca ccg 192
Pro Leu Pro Ser Pro Pro Pro Leu Ser Ala Pro Thr Ala Ser Pro Pro
50 55 60

```

```

cct ctt ccg gtt gaa tcc cca ccg tct cct cct ata gaa tca cca ccg 240

```


THE UNIVERSITY OF CHICAGO LIBRARY

gct Ala 305	acc Thr	aga Arg	aat Asn	aga Arg	aat Asn	aac Asn	aac Asn	agc Ser	agc Ser	tct Ser	gct Ala	cat His	cat His	caa Gln	tcc Ser	960
aaa Lys	act Thr	ccc Pro	tca Ser	aaa Lys	ggt Val	caa Gln	cat His	cat His	cgg Arg	ggc Gly	ggt Gly	aat Asn	gct Ala	ggt Gly	acg Thr	1008
aac Asn	cag Gln	gca Ala	cat His	ggt Val	atc Ile	aca Thr	atg Met	cca Pro	cca Pro	cca Pro	atc Ile	cat His	gct Ala	aaa Lys	tat Tyr	1056
ata Ile	tct Ser	agt Ser	gga Gly	ggt Gly	tgt Cys	gat Asp	acg Thr	aag Lys	gag Glu	aac Asn	aat Asn	tct Ser	ggt Val	gcg Ala	aaa Lys	1104
aac Asn	att Ile	tca Ser	atg Met	cca Pro	tct Ser	gga Gly	atg Met	ttc Phe	tcc Ser	tac Tyr	gaa Glu	gaa Glu	ctt Leu	tca Ser	aaa Lys	1152
gca Ala	act Thr	ggt Gly	gga Gly	ttt Phe	tca Ser	gag Glu	gag Glu	aac Asn	ctt Leu	ttg Leu	gga Gly	gaa Glu	ggc Gly	ggt Gly	ttc Phe	1200
gga Gly	tat Tyr	ggt Val	cac His	aaa Lys	gga Gly	gtg Val	ttg Leu	aaa Lys	aac Asn	ggg Gly	aca Thr	gaa Glu	ggt Val	gcg Ala	gtg Val	1248
aag Lys	cag Gln	ctg Leu	aag Lys	att Ile	ggg Gly	agc Ser	tat Tyr	caa Gln	ggg Gly	gaa Glu	aga Arg	gaa Glu	ttc Phe	caa Gln	gct Ala	1296
gag Glu	ggt Val	gac Asp	aca Thr	atc Ile	agt Ser	agg Arg	ggt Val	cat His	cat His	aag Lys	cac His	ctc Leu	ggt Val	tca Ser	ttg Leu	1344
ggt Val	ggt Gly	tat Tyr	tgc Cys	ggt Val	aat Asn	gga Gly	gat Asp	aaa Lys	aga Arg	ctc Leu	ttg Leu	ggt Val	tac Tyr	gag Glu	ttt Phe	1392
ggt Val	cct Pro	aaa Lys	gat Asp	acc Thr	ttg Leu	gag Glu	ttc Phe	cac His	ttg Leu	cat His	gag Glu	aac Asn	aga Arg	gga Gly	agc Ser	1440
gtg Val	ttg Leu	gaa Glu	tgg Trp	gaa Glu	atg Met	agg Arg	ctc Leu	agg Arg	att Ile	gct Ala	gta Val	gga Gly	gca Ala	gca Ala	aaa Lys	1488
gga Gly	tta Leu	gct Ala	tat Tyr	ctt Leu	cat His	gag Glu	gat Asp	tgc Cys	agt Ser	cca Pro	act Thr	ata Ile	att Ile	cac His	cgt Arg	1536
gat Asp	atc Ile	aaa Lys	gca Ala	gct Ala	aat Asn	atc Ile	ctt Leu	cta Leu	gat Asp	tcc Ser	aaa Lys	ttt Phe	gag Glu	gca Ala	aag Lys	1584
gtc	tct	gac	ttt	gga	cta	gcc	aag	ttt	ttc	tca	gac	acc	aat	tca	tca	1632

Val	Ser	Asp	Phe	Gly	Leu	Ala	Lys	Phe	Phe	Ser	Asp	Thr	Asn	Ser	Ser	
530						535					540					
ttc	act	cat	atc	tct	act	cga	gtg	gta	gga	act	ttc	gga	tac	atg	gct	1680
Phe	Thr	His	Ile	Ser	Thr	Arg	Val	Val	Gly	Thr	Phe	Gly	Tyr	Met	Ala	
545					550					555					560	
cca	gaa	tac	gcg	tcc	agt	ggg	aaa	gta	act	gat	aaa	tca	gat	gta	tat	1728
Pro	Glu	Tyr	Ala	Ser	Ser	Gly	Lys	Val	Thr	Asp	Lys	Ser	Asp	Val	Tyr	
				565					570					575		
tcc	ttt	ggg	gtc	gtg	ctt	cta	gaa	ctc	atc	act	gga	cgt	cca	tca	att	1776
Ser	Phe	Gly	Val	Val	Leu	Leu	Glu	Leu	Ile	Thr	Gly	Arg	Pro	Ser	Ile	
			580					585					590			
ttc	gcc	aaa	gat	tct	tcc	aca	aac	cag	agt	tta	gta	gac	tgg	gcg	agg	1824
Phe	Ala	Lys	Asp	Ser	Ser	Thr	Asn	Gln	Ser	Leu	Val	Asp	Trp	Ala	Arg	
		595					600					605				
cca	ttg	ctt	acg	aaa	gca	atc	tct	gga	gaa	agt	ttt	gac	ttt	ctt	gta	1872
Pro	Leu	Leu	Thr	Lys	Ala	Ile	Ser	Gly	Glu	Ser	Phe	Asp	Phe	Leu	Val	
	610					615					620					
gac	tca	agg	ttg	gag	aag	aat	tac	gat	aca	act	cag	atg	gca	aac	atg	1920
Asp	Ser	Arg	Leu	Glu	Lys	Asn	Tyr	Asp	Thr	Thr	Gln	Met	Ala	Asn	Met	
625					630					635					640	
gct	gct	tgt	gct	gct	gct	tgc	ata	cgc	caa	tca	gct	tgg	ctt	cgg	cct	1968
Ala	Ala	Cys	Ala	Ala	Ala	Cys	Ile	Arg	Gln	Ser	Ala	Trp	Leu	Arg	Pro	
				645					650					655		
aga	atg	agc	cag	gta	gta	cgt	gct	ctt	gaa	ggc	gag	gtg	gcc	ctg	aga	2016
Arg	Met	Ser	Gln	Val	Val	Arg	Ala	Leu	Glu	Gly	Glu	Val	Ala	Leu	Arg	
			660					665					670			
aag	gtc	gaa	gag	act	ggg	aat	agc	gtg	acc	tat	agc	tct	tct	gaa	aac	2064
Lys	Val	Glu	Glu	Thr	Gly	Asn	Ser	Val	Thr	Tyr	Ser	Ser	Ser	Glu	Asn	
		675				680						685				
ccg	aat	gac	atc	aca	cca	cgg	tat	gga	aca	aat	aag	agg	aga	ttc	gac	2112
Pro	Asn	Asp	Ile	Thr	Pro	Arg	Tyr	Gly	Thr	Asn	Lys	Arg	Arg	Phe	Asp	
	690					695					700					
aca	ggg	tca	agc	gat	ggg	tac	act	tca	gaa	tat	gga	gtt	aac	cct	tct	2160
Thr	Gly	Ser	Ser	Asp	Gly	Tyr	Thr	Ser	Glu	Tyr	Gly	Val	Asn	Pro	Ser	
705					710					715					720	
cag	tcg	agc	agt	gaa	cat	caa	cag	gtg	aat	act	tag					2196
Gln	Ser	Ser	Ser	Glu	His	Gln	Gln	Val	Asn	Thr						
				725					730							

<210> 17

<211> 731

<212> PRT

<213> Arabidopsis thaliana

<400> 17

21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 10

<213> Brassica napus

<220>

<221> PEPTIDE

<222> (1)..(9)

<223> X(1) =F/Y

<400> 19

Gly Gly Phe Gly Ile Val Xaa Lys Gly
1 5

<210> 20

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<220>

<221> misc_feature

<222> (1)..(22)

<223> N(1) =i N(2) =i N(3) =T/g N(4) =i N(5) =i N(6) =A/g N(7)
=A/g N(8) =T/c

<400> 20

aanatncnng ccatnccnaa nn

22

<210> 21

<211> 8

<212> PRT

<213> Brassica napus

<400> 21

Asp Phe Gly Met Ala Arg Ile Phe
1 5

<210> 22

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 22

taaccaacaa gagaca

16

<210> 23

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

```

<220>
<221> misc_feature
<222> (1)..(23)
<223> N(1) =g/A  N(2) =i    N(3) =g/A  N(4) =T/C  N(5) =i    N(6) =i    N(7)
=A/q  N(8) =I    N(9) =A/g  N(10) =A/g

```

```
<400> 23
annannttng cnannccnaa ntc                                     23
```

```
<210> 24
<211> 8
<212> PRT
<213> Arabidopsis thaliana
```

<400> 24
Asp Phe Gly Leu Ala Lys Leu Leu
1 5

```
<210> 25
<211> 18
<212> DNA
<213> Artificial Sequence
```

<220>
<223> Description of Artificial Sequence: primer

```
<400> 25
ccactcccaa ctttcaac
```

```
<210> 26
<211> 27
<212> DNA
<213> Artificial Sequence
```

<220>
<223> Description of Artificial Sequence: primer

```
<400> 26
ggaaaagcttg catgcctgca ggtcgac
```

```
<210> 27
<211> 30
<212> DNA
<213> Artificial Sequence
```

```
<220>
<223> Description of Artificial Sequence: primer
```

```
<400> 27
cgcgctgcagg taatacgact cactataggg 30
```

10	20	30	40	50	60
TTAACTCTCT	GGTCTCCGTG	TCTCCTCTCT	TCTCCTGCTG	CTTCCTTTTA	ACACTCTCTT
70	80	90	100	110	120
CATTTCGCTT	TTTGATTTAG	ATCCAAAGAA	GCAGACATCT	CCTCGGCGCC	GTCTCCGGGG
130	140	150	160	170	180
ACTGGTTCGC	CTCCATCTCC	ACCATCAAAC	TCCACAACCA	CCACTCCTCC	TCCAGCTTCC
190	200	210	220	230	240
GTCCTCTCTC	CCACCACACC	TTCTTCTCCT	CCGCGGCCAT	CCACTATTCC	GACATCTCCT
250	260	270	280	290	300
CCTCCTTCTT	CTCGCTCTAC	ACCTTCTGCT	CCTCCTCCAT	CTCCACCAAC	TCCATCTACG
310	320	330	340	350	360
CCGGGATCTC	CACCTCCTCT	TCCTCAGCCG	TCTCCACCCG	CTCCAACCTAC	GCCCGGATCT
370	380	390	400	410	420
CCACCCGCAC	CTGTTACTCG	TCCTACTCGA	AACCCTCCAC	CTTCAGTCCC	AGGACCACCG
430	440	450	460	470	480
TCCAATCCTT	CACGCGAAGG	AGGATCTCCT	CGACCTCCAT	CTTCTCCCTC	GCCGCGCTCT
490	500	510	520	530	540
CCTTCTTCCC	ACGGTTTATC	AACAGGAGTG	GTGGTGGGAA	TCGCCATCGG	AGGAGTCGCT
550	560	570	580	590	600
CTGCTTGTA	TAGTGACTCT	GATTTGTCTC	CTCTGTAGCA	AGAAACGACG	GAGAGACGAA
610	620	630	640	650	660
GAAGATGCTT	ACTATGTTCC	TCCGCCACCT	CCTCCTGGTC	CCAAAGCCGG	AGSACCTTAC
670	680	690	700	710	720
GGTGGACAGC	AGCAACAATG	GCGGCAACAA	AACGCAACAC	CACCGTCAGA	TCATGTCGTG
730	740	750	760	770	780
ACGTCACTAC	CACCACCAAC	TAAGGCTCCA	TCTCCACCAC	GGCAACCTCC	TCCACCTCCA
790	800	810	820	830	840
CCACCGCCTT	TCATGAGCAG	CAGCGGCGGC	TCCGACTACT	CGGACCGTCC	AGTTCTTCCT
850	860	870	880	890	900
CCACCGTCTC	CAGGGCTTGT	GTTAGGCTTC	TCCAAAAGCA	CTTTCACATA	CGAGGAGCTA
910	920	930	940	950	960
GCTAGAGCCA	CCAATGGTTT	CTCCGAGGCG	AACTTGTTAG	GACAAGGCGG	GTTCCGGTTAC
970	980	990	1000	1010	1020
GTGCACAAAG	GTGTGTTGCC	TAGTGGGAAA	GAAGTTGCTG	TGAAGCAGTT	GAAAGTTGCC
1030	1040	1050	1060	1070	1080
AGTGGTCAGG	GAGAGAGGGA	GTTTCAGGCA	GAGGTTGAGA	TCATCAGCAG	AGTTCACCAC
1090	1100	1110	1120	1130	1140
AGGCATCTGG	TGTCTCTTGT	TGGTTATTGC	ATCGCCGGTG	CCAAAAGATT	GCTTGTCTAT
1150	1160	1170	1180	1190	1200
GAGTTTGTTC	CTAACACAA	TCTCGAGCTT	CACCTCCATG	GCGAGGGACG	GCCTACAATG
1210	1220	1230	1240	1250	1260
GAATGGAGCA	CCAGATTGAA	GATTGCTCTT	GGATCTGCTA	AAGGACTTTC	TTATCTTCAT
1270	1280	1290	1300	1310	1320
GAAGATTGCA	ATCCTAAAT	CATTCAACGT	GATATCAAGG	CTTCAACAT	ATTGATAGAT
1330	1340	1350	1360	1370	1380
TTCAAGTTTG	AAGCTAAGGT	TGCTGATTTT	GGTCTTGCTA	AGATTGCTTC	TGATACAAAC
1390	1400	1410	1420	1430	1440
ACGCATGTAT	CAACACGTGT	GATGGGAACC	TTTGGGTACT	TGGCTCCGGA	ATACGCTGCA
1450	1460	1470	1480	1490	1500
AGCGGAAAGC	TCACGGAGAA	GTCTGACGTT	TTCTCATTG	GCGTTGTGCT	TTTGGAGCTC
1510	1520	1530	1540	1550	1560

Figure 1(d) (continued on next page)